FILOSOFÍA Y DERECHO



A la sombra de Hume

Un balance crítico del intento de la neuroética de fundamentar la moral

Daniel González Lagier

Marcial Pons

Colección

Filosofía y Derecho

José Juan Moreso Mateos
Jordi Ferrer Beltrán
Adrian Sgarbi
(dirs.)

A LA SOMBRA DE HUME

Un balance crítico del intento de la neuroética de fundamentar la moral

DANIEL GONZÁLEZ LAGIER

A LA SOMBRA DE HUME

Un balance crítico del intento de la neuroética de fundamentar la moral

Marcial Pons

MADRID | BARCELONA | BUENOS AIRES

2017

La colección *Filosofía y Derecho* publica aquellos trabajos que han superado una evaluación anónima realizada por especialistas en la materia, con arreglo a los estándares usuales en la comunidad académica internacional.

Los autores interesados en publicar en esta colección deberán enviar sus manuscritos en documento *Word* a la dirección de correo electrónico *manuscritos@filosofiayderecho.es*. Los datos personales del autor deben ser aportados en documento aparte y el manuscrito no debe contener ninguna referencia, directa o indirecta, que permita indentificar al autor.

En caso de ser aceptada la publicación del original, el autor deberá adaptarlo a los criterios de la colección, los cuales se pueden encontrar, en formato PDF, en la página web *www.filosofiayderecho.es*.



La colección Filosofía y Derecho ha sido distinguida en 2017 con el sello de calidad en edición académica CEA-APQ, con mención especial de internacionalidad, promovido por la Unión de Editoriales Universitarias Españolas (UNE) y avalado por la Agencia Nacional de Evaluación de la Calidad y Acreditación (ANECA) y la Fundación Española para la Ciencia y la Tecnología (FECYT).

www.marcialpons.es

ISBN: 978-84-9123-421-0
Depósito legal: M. 27.711-2017

Diseño de la portada: Manuel Estrada. Diseño Gráfico

Fotocomposición: Milésima Artes Gráficas

Impresión: Elecé Industria Gráfica, S.L.
Polígono El Nogal
Río Tiétar 24, 28110 Algete (Madrid)

MADRID, 2017



Para Sofía, por primera vez.
Para Julia, de nuevo.
Para Merche, como siempre.
Porque así —con primeros pasos, insistencias
y certezas— se tejen nuestras vidas.

ÍNDICE

Pág.

PRÓLOGO*

1. «Los científicos muestran que el juicio moral depende de las emociones». Así titula el periódico *El País,* de 22 de marzo de 2007, la noticia en la que se describen ciertos experimentos de Antonio DAMASIO y Marc HAUSER. El subtítulo da cuenta del argumento usado para llegar a esta conclusión: «Cierto daño cerebral lleva a *decisiones éticamente discutibles* aunque la razón está intacta» (la cursiva es mía). A un grupo de seis pacientes con daños cerebrales que afectaban a zonas del cerebro que regulan las emociones se les plantearon preguntas del siguiente tenor (las típicas preguntas sobre dilemas morales relacionadas con el valor de la vida): «Un amigo está infectado por un virus y plantea contagiar a otros. Algunos morirán. Tu única opción es dejar que ocurra o matarle. ¿Aprietas el gatillo?». El resultado es que, mientras la mayoría de la gente «normal» (sin los daños) responde que no, los pacientes de DAMASIO y HAUSER, según la noticia, «apretarían el gatillo». «Debido a su daño

---

* Esta investigación se ha desarrollado en el marco de los proyectos «Desarrollo de una concepción argumentativa del Derecho» (DER2013-42472-P) y «Prueba y atribución de responsabilidad: definición y contrastación del daño» (DER2014-52130-P) financiados por el Ministerio de Educación y Ciencia. La lista de las personas que han contribuido con sus comentarios a mejorar este trabajo es larga. Agradezco a mis compañeros del área de Filosofía del Derecho de la Universidad de Alicante todas sus observaciones, que sin duda han sido esenciales para permitirme verlo de una manera más objetiva. Victoria ROCA ha leido generosamente varias versiones del trabajo y me ha aportado sugerentes ideas, al igual que Sebastián FIGUEROA. Mi agradecimiento es muy especial para Manuel ATIENZA, Juan RUIZ MANERO, José Juan MORESO, Marina GASCÓN y María José AÑÓN, que han formádo parte del tribunal de cátedra al que este trabajo se presentó y lo han enriquecido con sus comentarios.

cerebral —explica un miembro del equipo investigador— tienen unas emociones sociales anormales, carecen de compasión y de empatía». HAUSER declara que su trabajo «aporta la primera evidencia causal del papel de las emociones en los juicios morales» (luego veremos que HAUSER cambió de opinión respecto de la función de las emociones). Esto se infiere del hecho de que los sujetos que no podían sentir emociones tomaban decisiones éticamente *discutibles*.

Esta noticia permite ver el tipo de método experimental usado por un buen grupo de neurocientíficos, médicos, psicólogos, etc. para estudiar el razonamiento y el comportamiento moral. También evidencia algunos de los problemas de estas investigaciones. La «neuroética» (la confluencia de la ética y el estudio del cerebro) pretende en algunos casos construir una ética científica, como en su día el «darwinismo moral» de SPENCER, que trataba de fundamentar las normas éticas en el proceso evolutivo, o la «sociobiología» de E. O. WILSON, que pretendía explicar el comportamiento de todos los animales extrapolando el egoísmo genético del mundo de los genes a los demás ámbitos de la vida (en realidad, la neuroética es una actualización de estas concepciones). Ello exige que sus postulados descansen en la observación y la experimentación. Para ello, los investigadores diseñan «experimentos morales», como el del ejemplo anterior y otros que veremos más adelante. Sin embargo, se debe ser muy cuidadoso con la interpretación de estos experimentos, con los presupuestos de tales interpretaciones y con los argumentos con los que se intenta fundamentar las conclusiones. Es un error pensar que los meros datos científicos pueden apuntar por sí solos a una u otras conclusiones. Aquí entran en juego cuestiones metodológicas, conceptuales y filosóficas que son imprescindibles para dar sentido a esos datos. Quienes estén preocupados por su salud mental por simpatizar con la respuesta consecuencialista o utilitarista en el ejemplo anterior (aunque quizá ellos mismos no dispararían, a pesar de creer que es lo correcto) pueden estar tranquilos, porque la noticia de *El País* revela serios errores «filosóficos». Por ejemplo, se generaliza a partir de solo seis casos; se pasa del «ser» (lo que los sujetos «normales» suelen decidir) al «deber ser» (lo que se debe decidir); se comete una petición de principio al asumir como correcta la solución que involucra a las emociones (a la otra se la considera «discutible») y tratar de derivar de ahí que las emociones

son necesarias para tomar las decisiones correctas; y se descarta sin fundamento la hipótesis alternativa de que la respuesta correcta sea la utilitarista y que las emociones sean un obstáculo que entorpece la decisión adecuada.

No quiero sugerir que este tipo de problemas vicie todas las afirmaciones de la neuroética. Para ser justos, hay que señalar que es muy posible que en el ejemplo anterior los defectos de razonamiento se deban a la simplificación hecha por el periodista. Pero esto nos lleva a otro problema preocupante: el impacto que estas investigaciones tienen en los medios de comunicación está contribuyendo a formar una imagen de la moral y de nuestra propia naturaleza a partir de bases todavía poco sólidas y de teorías muchas veces deformadas por el proceso de divulgación. La neuroética tiene mucho que aportar, pero hay que ser prudente frente a la espectacularidad de los aparentes resultados. Sobre todo porque una buena parte de la neuroética parece empeñada en enfocar solo el lado de la naturaleza humana precisamente menos peculiar de los humanos: empeñada, en definitiva, en negar la razón y las emociones más vinculadas a ella, en lugar de indagar sobre cómo son estas posibles. Para evitar este tipo de errores, interpretaciones sesgadas o precipitadas y falacias puede ser útil revisar algunos aspectos de la discusión sobre el origen de los juicios morales en la neuroética y tratar de sopesar su nivel de confirmación a partir de argumentos filosóficos. Este es el objetivo principal de este trabajo. Para ello, los dos primeros capítulos son predominantemente expositivos (postergando la mayoría de objeciones) y los dos últimos críticos y evaluativos.

2. La cuestión de cuál es el origen de nuestras opiniones morales —si la razón o las pasiones— ha sido también el objeto de un largo debate en filosofía, demasiado largo para poder entrar en él en un trabajo de este tipo. No obstante, creo que es importante tener ciertas coordenadas para poder evaluar la contribución de la neuroética, especialmente si pretende tener una dimensión normativa. De hecho, varios de los autores que veremos tratan de situarse a la sombra de algún importante filósofo (HUME, ARISTÓTELES, RAWLS), y la mayoría de ellos opinan que, respecto de la controversia acerca del origen de los juicios morales, HUME tenía razón. Por ello, en el capítulo I me ocupo brevemente de la teoría moral emotivista de HUME y de algunas de las teorías metaéticas del siglo XX que han tra-

tado de responder a esta cuestión (el intuicionismo y el emotivismo analítico). Veremos que la cuestión de si los juicios morales proceden o no de la razón se entrelaza con la cuestión de si las propiedades morales (aquellas propiedades de las acciones, las normas o los estados de cosas que nos hacen pensar que son moralmente buenos o malos) son propiedades *naturales* (empíricamente contrastables) o no, y esta, a su vez, con el problema de su objetividad. Me parece que este repaso nos permitirá mostrar (como veremos en el epílogo) que la teoría de la justicia que surge de los estudios de la neuroética es una extraña combinación de teorías metaéticas, y también nos permitirá plantear (en el capítulo IV) si realmente podemos derivar conclusiones normativas de la neuroética o si este «salto» a la normatividad, por el contrario, no está justificado.

En el capítulo II expongo algunas de las principales aportaciones (del lado de la neuroética y la psicología moral experimental, pues se trata de dos ámbitos que a veces son difíciles de distinguir) al problema de la relación entre las «corazonadas» y los juicios morales. Toda selección es un proceso de «poda» que moldea el tema que se estudia, y una poda inadecuada puede llegar a deformarlo. Es posible que la versión de la «neuroética» que se desprenda de mi selección no sea la única. Quizá, incluso, pueda considerarse injusta o pensarse que estoy incurriendo en la «falacia del hombre de paja», es decir, que estoy construyendo un «enemigo» a medida para poder combatirlo cómodamente. Como defensa puedo argüir que he optado por (algunos de) los autores que se han situado en el centro de la discusión neuroética, que han recibido mayor atención (y, criterio muy reconocido hoy día, mayor número de citas) y también que han tenido mayor presencia en trabajos de otras áreas e, incluso, en los medios de comunicación y redes sociales de divulgación científica (quizá precisamente por sostener tesis menos matizadas). Cuando me refiero a las aportaciones de la neuroética estoy pensando en la de estos autores, aunque hay que ser consciente de que hay muchos otros no representados en este trabajo y que no pueden caer bajo las mismas críticas.

Creo que estas aportaciones pueden situarse en dos niveles: explicativo y normativo, pero no siempre es fácil colocar a los autores en uno u otro nivel. He tratado de ordenarlos comenzando con los que solo tienen pretensiones descriptivas y explicativas y con-

tinuando con los que, de una manera crecientemente abierta, tratan de extraer conclusiones normativas y hasta proponen fundamentar una ética universal basada en el cerebro.

Me parece que las principales tesis descriptivas de la neuroética son las siguientes: 1) El juicio moral descansa fundamentalmente en las intuiciones; 2) las intuiciones morales las experimentamos como emociones de agrado o desagrado ante ciertas acciones o situaciones; 3) una parte relevante de la moral (o de nuestras capacidades morales) es innata; 4) el proceso de selección a través de la evolución puede explicar por qué tenemos unas intuiciones morales y no otras. En el capítulo III trato de mostrar que esas tesis deben tomarse como especulaciones sugerentes (y proyectos interesantes de investigación), pero no como conclusiones sólidas —al menos tal como son presentadas—, fundamentalmente porque las pruebas que se aportan a favor de ellas no logran excluir otras hipótesis alternativas que cambiarían la manera de entender la moral que tiene la neuroética (por ejemplo, dando mayor cabida a la razón o a la cultura, la sociedad y la educación). La historia de las ideas sugiere que a veces solo se llega a una visión ponderada sobre un tema tras una etapa más radicalizada. Creo que las aportaciones descriptivas de la neuroética pueden contener un importante grado de verdad, pero están aún en una «fase de exaltación» y deberán ser matizadas con ingredientes de la visión contraria (racionalista).

Por lo que respecta a la pretensión normativa, esto es, a la posibilidad de que el estudio de la moral desde una perspectiva neurocientífica (o científica, sin más) pueda decirnos qué teoría ética es —normativamente— más acertada o pueda servir para justificar principios o valores morales, en mi opinión no puede superar la clásica objeción (atribuida precisamente a HUME) de que es imposible fundamentar normas exclusivamente en hechos. El capítulo V está dedicado a mostrar que los intentos de autores como CHURCHLAND, HAUSER, CASEBEER y otros de eludir la *Ley de Hume* no son convincentes[1] y, por tanto, respecto de la discusión moral, la neuroé-

---

[1] Por el contrario, creo que el argumento debido a MOORE conocido como la falacia naturalista contiene una petición de principio, por lo que, aunque esto no demuestre que lo que MOORE quería concluir sea falso, no es un buen argumento contra las pretensiones de las éticas naturalistas.

tica solo puede ofrecer aportaciones a las premisas fácticas de los argumentos normativos, pero no puede por sí sola establecer una conexión válida entre nuestros principios morales últimos y la evolución o la estructura del cerebro. No obstante, como sugeriré en el Epílogo, quizá sí pueda contribuir a comprender el marco o ámbito en el que tienen sentido las exigencias de la moral[2].

En definitiva, parece que la sombra de HUME se extiende en dos direcciones: una —su emotivismo— favorable a las tesis de la neuroética; otra —la Ley o «guillotina» de Hume— contraria a ellas. En todo caso, es innegable que su obra ilumina gran parte de la discusión actual sobre las relaciones entre la ciencia y la ética.

[2] Las objeciones que planteo a las tesis descriptivas y normativas de la neuroética podrían, desde luego, ampliarse. Me limito a algunas de las que me han suscitado los trabajos analizados, pero sería posible plantearse otras de carácter más general, por ejemplo sobre la confusión entre *a)* encontrar el correlato cerebral de una función y *b)* explicación de dicha función o sobre el reduccionismo eliminacionista que inspira a las *neuro-disciplinas* (neuroética, neuroeconomía, neuroestética, neurorretórica...) (GARCÍA-ALBEA, 2011).

CAPÍTULO I

LA DISCUSIÓN EN LA FILOSOFÍA

1. LA PREGUNTA

¿Son nuestras opiniones morales fruto de nuestros razonamientos o, por el contrario, son el resultado de intuiciones o corazonadas? Una de las conclusiones que parece apoyar el tipo de investigaciones que se han englobado bajo el rótulo de «neuroética» es que más lo segundo que lo primero. Frente a las concepciones de la ética racionalistas, en la línea de KANT, la neuroética parece dar la razón a cierto intuicionismo moral (aunque muy distinto, como veremos, al intuicionismo filosófico de autores como G. E. MOORE, W. D. ROSS, H. A. PRICHARD, etc.) o al emotivismo moral (esto es, a concepciones de la ética que vinculan los juicios morales con las emociones y los deseos, en la línea de HUME).

La neuroética es una disciplina joven. Suele situarse su nacimiento en un congreso celebrado en San Francisco en el año 2002 dedicado a las relaciones entre la ética y la neurociencia (aunque la expresión «neuroética» ya se venía usando). A partir de ese momento han proliferado congresos, publicaciones, instituciones y cátedras sobre este tema. En realidad, la expresión «neuroética» sirve para dar cuenta de dos tipos de investigaciones, que es útil (aunque no siempre fácil) distinguir (ROSKIES, 2002: BONETE PERALES, 2010: 64; CORTINA, 2010: 131]: la *ética de la neurociencia*, una par-

te de la bioética que trataría de establecer un marco ético para las investigaciones neurocientíficas y sus aplicaciones; y la *neurociencia de la ética*, el estudio de la conducta ética desde el punto de vista de las investigaciones sobre el cerebro. Por ejemplo, son problemas de la ética de la neurociencia los siguientes: si está justificado o no el uso de los descubrimientos neurocientíficos para la mejora de las capacidades mentales o sensoriales de los humanos (el llamado «transhumanismo»); en qué condiciones es legítimo el uso en los tribunales de pruebas basadas en técnicas neurocientíficas (como la prueba P300 o *brainfingerprinting*, que permite determinar si el sujeto miente observando las variaciones en las ondas cerebrales ante ciertos estímulos); qué valor en relación con la atribución de responsabilidad hay que conceder a determinadas disfunciones cerebrales; o si es correcto —y en qué casos— usar técnicas de control de la conducta basadas en conocimientos neurocientíficos. Y son problemas de la neurociencia de la ética los siguientes: la discusión general sobre el libre albedrío y su relación con la responsabilidad; la cuestión de si nuestra actividad cerebral en el momento de tomar decisiones morales apunta más al deontologismo, al consecuencialismo o a una ética de las virtudes; el análisis del papel de la oxitocina o de las llamadas «neuronas espejo» en nuestra conducta ética; o si la neurociencia puede fundamentar conclusiones normativas acerca de la corrección de nuestros juicios morales. La cuestión que he planteado al comienzo y que discutiremos en estas páginas —si nuestros juicios morales proceden de la razón o de corazonadas— es, igualmente, un problema de la neurociencia de la ética.

2. EL EMOTIVISMO DE HUME

La afirmación de que nuestros juicios morales no provienen de la razón, sino de nuestras intuiciones y emociones, puede encontrarse en la filosofía de la moral ya antes del despliegue de la neurociencia. Las posturas de KANT y de HUME pueden tomarse como ejemplo de opiniones antagónicas sobre esta cuestión: HUME afirmó que «nuestro sentido del deber sigue en todo momento el curso común y natural de nuestras pasiones» (HUME, 1981: 708), que «la moralidad es, pues, más sentida que juzgada» (HUME, 1981: 691)

y que «el crimen o la inmoralidad no es un hecho particular o una relación, que puede ser objeto del entendimiento, sino que surge por entero del sentimiento de desaprobación, que, debido a la estructura de la naturaleza humana, inevitablemente experimentamos tras una aprehensión de la barbarie o la traición» (HUME, 1993: 204). Por el contrario, para Kant saber cuál es nuestro deber moral requiere un razonamiento por medio del cual hemos de comprobar si nuestros deseos se ajustan al imperativo categórico: «Para saber lo que he de hacer para que mi querer sea moralmente bueno [...] bástame preguntar: ¿puedes creer que tu máxima se convierta en ley universal? Si no, es una máxima reprobable, y no por algún perjuicio que pueda ocasionarte a ti o a algún otro, sino porque no puede convenir, como principio, en una legislación universal posible» (KANT, 1973: 43). Por ejemplo, «para resolver de la manera más breve la pregunta de si una promesa mentirosa es conforme al deber, me bastará preguntarme a mí mismo: ¿me daría yo por satisfecho si mi máxima —salir de apuros por medio de una promesa mentirosa— debiese valer como ley universal tanto para mí como para los demás? ¿Podría yo decirme a mí mismo: cada cual puede hacer una promesa falsa cuando se halla en un apuro del que no puede salir de otro modo? Y bien pronto me convenzo de que, si bien puedo querer la mentira, no puedo querer, empero, una ley universal de mentir; pues, según esta ley, no habría propiamente ninguna promesa, porque sería vano fingir a otros mi voluntad respecto de mis acciones futuras, pues no creerían ese mi fingimiento, o si, por precipitación lo hicieren, pagaríanme con la misma moneda; por tanto, mi máxima, tan pronto como se tornase ley universal, destruiríase a sí misma» (KANT, 1973: 42-43). De manera que «no puede ser admitido ninguna especie particular de sentimiento con el nombre de práctico o de moral, como sentimiento que precediese a la ley moral y estuviera a su base» (KANT, 1996: 133).

Lo anterior no quiere decir que Kant restara toda importancia al sentimiento moral [que, en algún pasaje, identifica con el «respeto a la ley», aunque es un sentimiento provocado por la razón (KANT, 1996: 134)] o que HUME no vea el espacio de la razón en la moral (que se advierte en su recurso al principio de utilidad), pero sí muestra el énfasis que, a juicio de estos autores, tenía uno u otro elemento en la generación de los juicios morales.

Vale la pena que nos detengamos en la filosofía moral de HUME, porque en bastantes aspectos enlaza con algunas de las afirmaciones actuales de la neuroética.

El emotivismo de HUME no implica que para este autor la moral sea un fenómeno plenamente natural. Por el contrario, para HUME la justicia es una virtud artificial, aunque con una base natural, aprendida por el hombre a través de convenciones: «El sentido de la justicia y de la injusticia —escribe— no se deriva de la naturaleza, sino que surge, de un modo artificial aunque necesario, de la educación y de las convenciones humanas» (HUME, 1981: 707). Podríamos reconstruir el proceso de «invención» de la justicia de la siguiente manera. En un primer momento, «es natural que el hombre quiera a sus hijos más que a sus primos, a sus primos más que a los extraños, con los que todo lo demás da igual». Es de estas preferencias de donde surge nuestro sentido del deber. Sin embargo, en un segundo momento nuestro propio interés en sobrevivir en un mundo de escasez y en unas circunstancias de vulnerabilidad nos empuja a unirnos en sociedad (en círculos, por tanto, más amplios) y establecer las reglas de la justicia, de manera que lo que es bueno (útil) para la sociedad, es bueno para sus integrantes, y el ataque a las bases de la sociedad, lo es a sus integrantes. La sociedad es la solución a la debilidad humana[1]. El origen de la justicia, por tanto, no es primariamente una benevolencia general o el amor a la humanidad, sino

---

[1] «De todos los animales que pueblan el globo, no existe otro con quien la naturaleza haya parecido ser más cruel, a primera vista, que con el hombre, dadas las innumerables carencias y necesidades de que la naturaleza le ha provisto y los limitados medios que le proporciona para la satisfacción de estas necesidades. Por lo general, en otras criaturas se compensan estos extremos: si juzgamos que el león es un animal voraz y carnívoro, fácilmente descubriremos que tiene muchas necesidades, pero si atendemos a su figura y carácter, a su agilidad, valor, garras y fuerza, veremos que estas ventajas están en proporción con lo que le falta. La oveja y el buey carecen de esas ventajas, pero sus apetitos son moderados y su alimento fácil de conseguir. Solo en el hombre es posible observar en forma extrema esa conjunción antinatural de necesidad y debilidad. No solo el alimento que necesita para sustentarse escapa cuando lo busca y cuando se acerca a él, o por lo menos le exige grandes esfuerzos, sino que necesita también de vestidos y albergues para defenderse de la intemperie. Y, sin embargo, si consideramos al hombre solamente en sí mismo, vemos que no está provisto de garras y que no tiene fuerza ni ninguna otra capacidad natural que pudiera corresponder de algún modo a tantas necesidades como tiene. Solo reuniéndose en sociedad es capaz de suplir sus defectos y llegar a ser igual que las demás criaturas, y aun de adquirir superioridad sobre ellas» (HUME, 1981: 709).

el autointerés: «Una vez que los hombres han visto por experiencia que su egoísmo y su limitada generosidad los incapacitan totalmente para vivir en sociedad, si esas pasiones actúan a su arbitrio, y han observado al mismo tiempo que la sociedad es necesaria para satisfacer esas mismas pasiones, se ven naturalmente inducidos a someterse a la restricción de tales reglas» (HUME, 1981: 725-726). Por tanto, «el origen de la justicia se encuentra únicamente en el egoísmo y la limitada generosidad de los hombres, junto con la escasa provisión con que la naturaleza ha subvenido a las necesidades de estos» (HUME, 1981: 722). Ahora bien, en un tercer momento, de ese interés egoísta surge la *simpatía* hacia los demás. La simpatía, para HUME, al menos en el *Tratado*, no es la *benevolencia* (esto es, la *buena voluntad* hacia las personas), aunque puede explicarla, sino que es más bien lo que hoy llamaríamos empatía (en sentido estricto): es más bien (por decirlo con PÉREZ BERMEJO) «un proceso psicológico de identificación con los sentimientos ajenos, un mecanismo que supone en cada ser humano la facultad de autorrepresentación de las emociones de placer o de dolor del resto de los seres» (PÉREZ BERMEJO, 1999: 92). Y en palabras de HUME: «Cuando percibo los efectos de la pasión en la voz y el gesto de una persona, mi mente pasa de inmediato de estos efectos a sus causas, y se hace una idea tan vivaz de la pasión que al instante la convierte en esa misma pasión. De igual modo, cuando me doy cuenta de las causas de una emoción, mi mente pasa a los efectos que esas causas producen, con lo que se ve movida por una emoción similar [...]. Por estas cosas es por lo que inferimos la pasión y son ellas, en consecuencia, las que dan origen a nuestra simpatía» (HUME, 1981: 821). La simpatía es fundamental, en HUME, para entender el origen de los juicios morales. Mientras que el interés es la fuente de la *obligación natural*, la simpatía es la fuente de la *obligación moral*, y permite explicar que «aun cuando la injusticia esté tan alejada de nosotros que no afecte en modo alguno a nuestros intereses, nos sigue disgustando porque la consideramos nociva para la sociedad humana y perniciosa para todo el que se acerque al culpable. Participamos *por simpatía* del malestar del afectado» (HUME, 1981: 726). En esto, por cierto, HUME (al igual que otros autores, como Adam SMITH y su común maestro Francis HUTCHESON) se adelanta a las teorías neuroéticas que hoy día enfatizan la importancia de las llamadas «neuronas espejo» en la conducta moral. Se trata de un tipo de neuronas, descu-

biertas en 1996 por Giacomo Rizzolatti en los simios, incluyendo los humanos (aunque de ello solo hay pruebas indirectas), que se activan tanto cuando realizamos una acción como cuando vemos que otros la realizan, y que se supone que posibilitan, sin necesidad de razonamientos o inferencias, comprender las acciones del otro, sus intenciones y sus emociones (Rizzolatti; Craighero, 2004). Serían, por tanto, las neuronas de la «empatía» (o la «simpatía» de Hume) (para una comparación entre la noción de empatía en Adam Smith y en la neurociencia Shand, 2014) y, de acuerdo con esta hipótesis, tendrían un importante papel en la vida social y moral de los humanos (Iacoboni, 2013).

¿Es la filosofía moral de Hume una propuesta «naturalista»? Situar a Hume en las clasificaciones más habituales de las teorías éticas es un asunto discutido y complejo. Responder a la anterior pregunta depende de qué se entienda por «naturalismo ético». Si el naturalismo ético es un tipo de realismo moral que sostiene que las propiedades morales son propiedades (naturales: esto es, observables, empíricas) *de las cosas de las que decimos que son buenas, justas, etc.*, entonces es difícil considerar la filosofía moral de Hume como naturalista, puesto que para él los valores morales no son algo que esté en las cosas o en las acciones (no son hechos ni relaciones: si lo fueran, los juicios morales dependerían de la razón, no de la pasión), sino algo que proyectamos sobre ellas (en ocasiones Hume parece considerarlas cualidades secundarias o subjetivas, en el sentido de Locke). Recuérdese que, para Hume, no podemos pasar del «ser» (la descripción de las propiedades naturales de las cosas) al «deber ser», como hacen las éticas naturalistas en este sentido. Si por «naturalismo ético» entendemos el intento de explicar la moral desde un punto de vista empírico y de reducir los hechos morales a hechos no-morales, entonces sí lo es: el empirismo de Hume, que trata de dar una explicación de la moral desde datos empíricos, como la vulnerabilidad humana, sus pasiones y sus circunstancias, permite afirmar que su filosofía moral es naturalista en este sentido. Por último, el propio Hume afirmaba que, aunque la justicia es una virtud artificial y, por tanto, no natural, en otro sentido (un tercer sentido) sí se puede considerar natural: las reglas de la justicia son naturales en el sentido de que no son arbitrarias, sino necesarias, inseparables de la naturaleza y las circunstancias humanas: «La

humanidad es una especie inventiva; y cuando una invención es obvia y absolutamente necesaria puede decirse con propiedad que es natural» (HUME, 1981: 708). Con este comentario HUME parecería estar distanciándose del escepticismo y relativismo moral y tratar de retener un mínimo de contenido moral común: la necesidad de vivir en sociedad, la necesidad de la justicia e incluso su contenido (al menos, sus principios básicos) no pueden obviarse. En ese caso, el emotivismo de HUME sería distinto del emotivismo analítico que veremos más adelante.

3. EL INTUICIONISMO MORAL

Aunque la neuroética actual, como veremos, parece entrelazar las intuiciones con las emociones, lo que en filosofía se ha llamado emotivismo moral no debe ser confundido con lo que se ha llamado intuicionismo moral. Hemos visto que algunas teorías éticas consideran que términos como «bueno», «justo», etc., hacen referencia a propiedades naturales, de manera que son definibles por referencia a estas. A esta postura se le ha llamado naturalismo ético. G. E. M. MOORE criticó duramente a estas teorías, con su conocido argumento de la «falacia naturalista» (de la que la imposibilidad de pasar del «ser» al «deber ser» señalada por HUME sería un caso). En su opinión, todo intento de definir las propiedades morales en términos de propiedades naturales incurre en esta falacia, que surge del error de pensar que propiedades coexistentes son la misma propiedad: una cosa «buena» puede ser también útil, deseada, placentera, etc., pero eso no quiere decir que «bueno» y «útil», o «bueno» y «deseado» («placentero», etc.) sean lo mismo. Que dos cosas vayan siempre juntas no quiere decir que sean idénticas. Por ello, si afirmamos que «lo bueno es la felicidad», siempre tiene sentido preguntarse por qué la felicidad es buena (cosa que no tendría sentido si fueran idénticas). Para MOORE «bueno» es un término indefinible, una cualidad simple (como «amarillo»), porque no puede descomponerse (al contrario, por ejemplo, que «caballo»: «puede darse una definición de un caballo —escribe— porque *un caballo* tiene múltiples propiedades diferentes, todas las cuales pueden enumerarse. Pero una vez enumeradas, cuando se haya reducido el caballo a sus

más simples términos, entonces, no se puede ya más definir estos» (MOORE 1959: 6).

MOORE, por tanto, no niega que existan propiedades morales; lo que niega es que se trate de propiedades naturales. Pero entonces surge la cuestión de cómo pueden ser conocidas tales propiedades. MOORE es —como los naturalistas éticos— un realista moral. Pero si el método apropiado para captar o examinar las propiedades naturales no sirve para las propiedades morales, ¿cómo podemos tener acceso a las propiedades morales? Su respuesta y la de otros autores, como H. A. PRICHARD y W. D. ROSS, es que conocemos lo que es bueno (recto, justo, etc.) por intuición. El intuicionismo moral, por tanto, es un tipo de realismo moral, pero difiere del naturalismo en que niega que las propiedades morales sean propiedades naturales (son, sencillamente, propiedades de un tipo distinto: propiedades morales) y en que el modo de conocer tales propiedades dependa de la razón y el método científico.

¿En qué consiste la intuición moral? Los intuicionistas no fueron muy claros al respecto (en realidad, no parece que sea posible definir con claridad qué es la intuición, aunque, al mismo tiempo, hay que reconocer que muchas veces decimos tener intuiciones; más adelante volveremos sobre ellas). Tampoco es seguro que entendieran lo mismo por «intuición». Por ejemplo, entre PRICHARD y ROSS, por un lado, y MOORE, por otro, habría —al menos— tres diferencias respecto de la manera de entender las intuiciones morales: en primer lugar, PRICHARD las define «como una aprehensión inmediata exactamente en el sentido en que una aprehensión matemática es inmediata, por ejemplo, la aprehensión de que esta figura de tres lados, por tener tres lados, debe tener también tres ángulos. Ambas aprehensiones son inmediatas, en el sentido de que, en ambas, la penetración en la naturaleza del objeto conduce directamente al reconocimiento de que posee el predicado en cuestión» (PRICHARD, 1999: 8). También ROSS señala que las obligaciones *prima facie* «llegan a sernos evidentes de suyo exactamente como lo hacen los axiomas matemáticos» (ROSS, 1994: 48) y da un catálogo de las mismas (que, paradójicamente, pese a su supuesta evidencia, considera solo aproximado). Sin embargo, para MOORE afirmar que una proposición es una «intuición» solo quería decir que no admite prueba, sin querer hacer referencia a un modo o capacidad espe-

cial de conocimiento (HUDSON, 1987: 87). En segundo lugar, para PRICHARD y ROSS el objeto de nuestras intuiciones morales son las obligaciones o deberes (son intuicionistas deontológicos), mientras que para MOORE lo son los fines valiosos o buenos (MOORE combina el intuicionismo y el teleologismo: las intuiciones nos dicen qué fines son buenos y las acciones son justas en la medida en que son un medio para esos fines. De acuerdo con MOORE, es evidente que las cosas más valiosas son el disfrute de la buena compañía y los placeres de las cosas bellas (HUDSON, 1987: 92). En tercer lugar, para PRICHARD y ROSS las intuiciones de las personas suficientemente informadas son infalibles, mientras que para Moore es posible tener una intuición falsa (HUDSON, 1987: 88).

Como es imaginable, el punto más controvertido del intuicionismo es esta difusa facultad o este carácter evidente de lo «bueno». ¿No es posible tener intuiciones equivocadas? ¿No puede ocurrir que las personas tengan intuiciones morales distintas sobre el mismo asunto? PRICHARD y ROSS piensan que las personas pueden tener grados distintos de desarrollo moral y tener por ello intuiciones más o menos claras, lo que podría explicar las equivocaciones y los aparentes conflictos; también aceptan que junto con la intuición entra en juego el razonamiento para el examen de la situación que se está evaluando (para saber qué es lo justo en una determinada situación primero hay que analizarla, darse cuenta de las circunstancias, antes de que la intuición nos dé la respuesta). Sin embargo, la principal crítica es que la idea de intuición parece ser vacía, carente de contenido. De acuerdo con la definición habitual de «conocimiento», para poder decir que conocemos o sabemos algo (x) deben darse tres condiciones: 1) Que x sea verdad, 2) que creamos que x y 3) que tengamos buenas razones para creer x. Sin embargo, la idea de intuición no parece adecuada para satisfacer la tercera de estas condiciones (parece más una manera de eludir el problema de la justificación de nuestra creencia) y es redundante con la segunda: decir que sabemos x por intuición no parece que pueda diferenciarse de la insistencia en que creemos firmemente x. En palabras de HUDSON: «¿Qué nos dice "Sé por intuición que X" que no nos diga "creo que X"? La respuesta parece ser: nada» (HUDSON, 1987: 107).

4. EL EMOTIVISMO ANALÍTICO

Si se aceptan las anteriores críticas, tanto el realismo naturalista como el realismo intuicionista, las «dos vías del realismo moral» (LUQUE, 2011) parecen más bien dos callejones sin salida. El naturalismo ético lleva a la falacia naturalista; el intuicionismo ético, a postular una facultad misteriosa de conocimiento, lo que parece más bien una renuncia a tratar de justificar el conocimiento moral. Esto dio pie a un resurgir del emotivismo (en un sentido más radical que el de HUME), en buena parte motivado también por el positivismo lógico y su teoría verificacionista del lenguaje.

Como es sabido, para el positivismo lógico una proposición tiene significado solo si es verificable, esto es, si puede, por medios empíricos, confirmarse como verdadera o falsa (o bien si es una proposición analítica, verdadera o falsa por razones lógicas). Las proposiciones no analíticas que no pueden verificarse empíricamente no tienen sentido (recuérdese que, hasta el segundo WITTGENSTEIN y la denuncia de AUSTIN de lo que él llamaba la «falacia descriptivista», el significado de una oración solía identificarse exclusivamente con su sentido y referencia, y no con su uso, lo que implicaba que solo se pudiera hablar de significado descriptivo). A partir de aquí, algunos autores (B. RUSSELL, J. AYER y C. L. STEVENSON, entre otros, con antecedentes en el WITTGENSTEIN del *Tractatus* y en el pragmatismo ético de John DEWEY) afirmaron que los juicos de valor no son juicios sobre hechos y no tienen significado, sino que se limitan a expresar emociones o sentimientos. Con eso no quiere decirse que sean juicios que describen o informan de las emociones de los hablantes (en esto se distinguiría el *emotivismo* de estos autores del *subjetivismo*), sino una manera de expresarlas (como «¡Hurra!», «¡Ay!» o «¡Qué horror!»). En palabras de AYER, «en todos los casos en los que podría decirse que alguien está haciendo un juicio ético, la función de la palabra ética es puramente "emotiva". Es habitual expresar sentimientos acerca de determinados objetos, pero no hacer ninguna afirmación acerca de ellos» (AYER, 1984: 131). Sin embargo, junto con esta función expresiva, los emotivistas suelen admitir que los juicios de valor tienen también la función práctica de tratar de influir en el oyente y suscitar en él sentimientos que

le muevan a la acción. Esta función era el principal rasgo de los juicios de valor según STEVENSON. Para este autor, un juicio del tipo «esto es bueno» debe entenderse como «apruebo esto; apruébalo tú también» (donde «apruebo esto» debe entenderse como expresión, y no como descripción, de una emoción) (HUDSON, 1987: 120). Por medio de los juicios de valor y del uso de expresiones éticas (deber, correcto, etc.) lo que hacemos fundamentalmente es intentar influir en los demás. De una manera similar al prescriptivismo de HARE (para quien los juicios morales pertenecen al lenguaje directivo o imperativo), STEVENSON afirma que los juicos éticos pueden tener «una fuerza cuasi imperativa» (STEVENSON, 1993: 274)[2].

También un autor como G. H. VON WRIGHT ha defendido, más recientemente, el emotivismo como concepción de los valores o, como diría él, más exactamente de *las valoraciones*, tratando de establecer nuevas bases para esta teoría. En «Valorar (o cómo hablar de lo que se debe callar). Nuevas bases para el emotivismo» (2001) VON WRIGHT propone que los valores son ficciones lingüísticas y que las valoraciones tienen un fundamento emotivo (a diferencia de las normas, que tienen un fundamento volitivo) y no describen nada[3]. Y en el ámbito de la filosofía del derecho, uno de los más radicales defensores del emotivismo ético ha sido

---

[2] Dado el énfasis que STEVENSON pone en la función de «influencia» de los juicios morales, junto con algunos otros aspectos de su teoría, puede dudarse de que sea encuadrable en el emotivismo ético (a pesar de que suele citarse como el representante más típico de esta teoría) (ALARCÓN CABRERA, 2001: 155).

[3] Creo que de este trabajo cabe destacar los siguientes puntos: 1) «Una valoración es una actitud aprobatoria o desaprobatoria de un sujeto *s* respecto a un objeto *o*» (p. 387), de manera que «las valoraciones son actitudes emocionales» (p. 388), «no son ni verdaderas ni falsas» (p. 388) y son relativas a ciertos criterios y variables con el tiempo. 2) Es fundamental la distinción (no siempre tenida en cuenta) entre los valores propiamente dichos («bueno-malo», «bello-feo», «justo-injusto») y las cosas valoradas o bienes (democracia, libertad de expresión, igualdad de oportunidades). 3) Los juicios de valor (del tipo «Esto es bueno» o «me gusta esto»), cuando se usan evaluativamente (y no para informar de que el sujeto valora algo de cierta manera), no describen nada, sino que solo expresan las valoraciones del sujeto. 4) Mientras las normas se justifican «teleológicamente», las valoraciones se justifican (o explican) mediante referencias «a la forma de ser y al estilo de vida de los sujetos que valoran» (p. 391). 5) Los valores tienen un fundamento emotivo (expresan emociones) y las normas un fundamento volitivo (expresan la voluntad de la autoridad). 6) En los valores compartidos más importantes de una sociedad «hay un embrión de orden jurídico. Un orden jurídico tiene con frecuencia su origen en el hecho de que las valoraciones "cristalizan" en normas» (p. 392). 7) Los valores no son entidades, son más bien «ficciones lingüísticas» (p. 395).

probablemente Alf Ross, con su famosa frase: «Invocar la justicia es la misma cosa que golpear con el puño una mesa: una expresión emotiva que transforma una exigencia en un postulado absoluto» (A. Ross, 1963: 259).

Una de las principales críticas a las que ha tenido que enfrentarse el emotivismo es la de que deja a la ética fuera de la argumentación moral. Si los juicios morales son meras expresiones de emociones de agrado o desagrado, no parece que pueda haber discusión sobre ellos, como no discutimos sobre gustos. Si a mí me parece que el aborto puede estar en ocasiones justificado y a otro sujeto le parece que el aborto nunca lo está, y al decir esto solo estamos expresando nuestras emociones, entonces los dos juicios no entran en conflicto, no hay genuino desacuerdo. Y sin desacuerdo, no tiene sentido la argumentación moral. La respuesta de J. Ayer fue que las disputas morales no versan realmente sobre los valores en juego, sino sobre los hechos que valoramos. En palabras suyas: «Cuando alguien discrepa de nosotros acerca del valor moral de cierta acción o clase de acción, generalmente recurrimos al razonamiento, a fin de ganarle para nuestro modo de pensar. Pero no intentamos demostrar mediante nuestros argumentos que él tiene el sentimiento ético "injusto" respecto de una situación cuya naturaleza ha captado correctamente. Lo que tratamos de demostrar es que está equivocado acerca de los hechos del caso. Argüimos que ha interpretado mal los motivos del agente; o que ha juzgado mal los efectos de la acción [...]; o que no ha alcanzado a tomar en consideración las especiales circunstancias en que el agente se encontraba» (Ayer, 1984: 135). Y pensamos que es importante convencer de cuáles son los hechos porque confiamos en que el otro sujeto tenga una educación moral semejante y reaccione igual que nosotros ante los mismos hechos. La respuesta de Stevenson, como es sabido, consiste en distinguir entre desacuerdos de creencias (acerca de los hechos del caso, o de posibles efectos de una u otra acción) y desacuerdos de actitud (la valoración propiamente dicha) y señalar que la argumentación solo puede consistir en tratar de influir causal o psicológicamente (pero no racionalmente) en las diferentes actitudes.

5. ALGUNAS CONCLUSIONES DE LA DISCUSIÓN EN LA FILOSOFÍA

En las anteriores páginas hemos revisado dos concepciones de la ética que impugnan la conexión entre los juicios morales y la razón, vinculándolos con emociones e intuiciones: el emotivismo moral, en su versión humeana y en su versión analítica, y el intuicionismo ético. Ambas son, sin embargo, bien distintas.

Como sabemos, las teorías éticas suelen clasificarse en torno a varios ejes (RIVERA LÓPEZ, 2002: 48):

1) Son descriptivistas (o realistas) si consideran que los juicios morales son descriptivos y, por tanto, verdaderos o falsos; en caso contrario, son teorías no descriptivistas.
2) Las teorías descriptivistas son naturalistas si creen que los hechos descritos por los juicios morales son hechos o propiedades naturales; son intuicionistas si creen que se trata de un tipo de hechos distinto, *sui generis*, que conocemos por intuición.
3) Las teorías naturalistas son subjetivistas cuando consideran que las entidades que describen los juicios morales tienen un carácter subjetivo (dependiente del sujeto que evalúa: preferencias, emociones, creencias, etc.); y son objetivistas cuando remiten a hechos objetivos (como necesidades básicas, estados físicos, etc.).
4) Las teorías no descriptivistas pueden ser emotivistas (los juicios morales expresan actitudes emocionales) o prescriptivistas (los juicios morales son imperativos o prescripciones universalizables).

Pues bien, mientras el intuicionismo es una ética descriptivista que considera que existen propiedades morales (aunque no cognoscibles por la razón), el emotivismo es una teoría no descriptivista (aunque en la versión de HUME, sin embargo, no tiene las mismas implicaciones subjetivistas que en la de STEVENSON o AYER). Además, las «intuiciones» de MOORE, ROSS o PRICHARD poco o nada tienen que ver con las emociones. El intuicionismo ético no hace descansar esa facultad en las emociones, aunque tampoco es un razonamiento o reflexión. Sin embargo, ambas vías, la intuicionista y la emotivista, no dan cabida plena a la discusión moral: esta, para el naturalismo, es una discusión racional que trata de probar que las

cosas (acciones, estados, etc.) evaluadas tienen o no cierta propiedad moral. Por el contrario, para el intuicionismo, la discusión moral queda reducida a la discusión acerca del análisis de los hechos del caso que se juzga, con la confianza en que, estando de acuerdo en los hechos, surjan las mismas intuiciones entre los antagonistas. Y, de manera semejante, para el emotivismo los desacuerdos pueden tratar sobre los hechos del caso o (como sugiere STEVENSON) sobre las actitudes que esos hechos suscitan. Si el desacuerdo se debe solo a diferentes actitudes, pero hay acuerdo sobre los hechos, solo cabe persuadir o influir psicológicamente para tratar de modificar la actitud.

Obviamente, ninguno de los autores que hemos visto hasta ahora conoce los recientes avances neurocientíficos (aunque VON WRIGHT estuvo atento a lo que la neurociencia podía aportar al problema de la relación mente-cuerpo). ¿Qué dice la neuroética sobre el origen de los juicios morales? ¿Puede la neurociencia y la psicología contribuir a esclarecer el problema? En lo que sigue me ocuparé de algunos autores cuyos trabajos suelen ser considerados entre los textos básicos de la neuroética a propósito del problema del conocimiento moral.

CAPÍTULO II

LA DISCUSIÓN EN LA NEUROÉTICA

1. EL INTUICIONISMO SOCIAL DE J. HAIDT

En su trabajo «El perro emocional y su cola racional: un enfoque intuicionista social del juicio moral», Jonathan HAIDT propone un modelo de explicación del origen de los juicios morales al que denomina «intuicionismo social»: intuicionismo por el papel que concede a las intuiciones y emociones, y social por el papel que concede a la influencia de los aspectos sociales y culturales. Se contrapone, por tanto, al modelo racionalista imperante en las explicaciones del origen de las opiniones morales en psicología moral, que pone el acento en el razonamiento privado del individuo (como los modelos de KOHLBERG o PIAGET).

La tesis principal del intuicionismo social es, según el propio HAIDT, «que el juicio moral es causado por intuiciones morales fugaces y es seguido (cuando es necesario) por razonamientos morales lentos y *ex post facto*» (HAIDT, 2012: 168). El autor ilustra su tesis con un experimento y la justifica con varios indicios de la ineficacia causal de la razón. El experimento es el siguiente:

> Julia y Marcos son hermanos. Se encuentran de viaje en Francia durante las vacaciones de verano del colegio. Una noche se quedan solos en una cabaña junto a la playa. Consideran que sería interesante y divertido si intentaran hacer el amor. Al menos sería una nueva

> experiencia para cada uno de ellos. Julia estaba tomando ya píldoras anticonceptivas, pero Marcos usa un preservativo también por seguridad. Los dos disfrutan haciendo el amor, pero deciden no hacerlo más. Ambos guardan esa noche como un secreto especial que les hace sentirse más unidos. ¿Qué opinas de esto? ¿Fue correcto que hicieran el amor?
>
> Muchas personas que escuchan la historia anterior inmediatamente afirman que fue incorrecto que hicieran el amor, y a continuación comienzan a buscar razones que inciden en los riesgos de la endogamia, aun cuando Julia y Marcos emplearon dos tipos de anticonceptivos. Argumentan que Julia y Marcos resultarían dañados, quizás emocionalmente, incluso cuando la historia deja claro que no sufrieron ningún perjuicio. Ocasionalmente mucha gente dice cosas como «no lo sé, no puedo explicarlo, solo sé que es incorrecto» (HAIDT, 2012: 159).

HAIDT cree que este tipo de situaciones (este «desconcierto» al «sentir» que algo está mal pero no poder justificar por qué) es característico de nuestras actitudes morales y que una teoría que explique el juicio moral debe dar cuenta de estos casos. En su opinión, las teorías intuicionistas están mejor situadas para explicarlos: «El intuicionismo en filosofía —sostiene— se refiere a la concepción de que hay verdades morales y que cuando las personas comprenden estas verdades no lo hacen por un proceso de racionalización y reflexión sino más bien por otro más parecido a la percepción, en el que uno solo ve sin argumentos que aquellas son y deben ser ciertas» (HAIDT, 2012: 10). Los juicios morales son, entonces, causados directamente por estas intuiciones, sin que la razón tenga aquí papel causal alguno.

Los indicios que presenta HAIDT para dudar de que la razón tenga habitualmente un papel relevante en la formación de los juicios morales son los siguientes:

1. En primer lugar, acude al creciente consenso entre los psicólogos sobre el hecho de que cuando los sujetos están tratando de resolver un problema (y los problemas morales no serían distintos del resto en este punto) no solo se activa el proceso de razonamiento, sino que, simultáneamente, se activa otro proceso mucho más rápido que nos avanza la solución. Se trata del «modelo de procesamiento dual». Ahora bien, HAIDT cree que hay evidencias cientí-

ficas de que, de los dos procesos, el intuitivo y emocional es el que marca la pauta en la resolución de problemas, básicamente a través de una evaluación automática de las situaciones y las personas. Por ejemplo, respecto de la formación de las opiniones sobre otras personas (incluidos juicios sobre su carácter o sobre la corrección moral de las mismas o de sus acciones), las pruebas muestran que «las personas forman sus primeras impresiones a simple vista ( ), y las impresiones que se forman con observar "finas porciones" de comportamiento (tan escasas como cinco segundos) son casi idénticas a las que se forman mediante una mucho más larga y más relajada observación y deliberación» (Haidt, 2012: 174). Posteriormente, estas primeras impresiones determinan el resto de opiniones morales sobre esas personas.

2. En segundo lugar, algunas investigaciones avalan que ambos procesos no se dan simultáneamente, sino que el papel de la razón suele limitarse a justificar o dar sentido, *a posteriori,* a la opinión intuitivamente emitida. Y no solo es un razonamiento *post hoc,* sino que, además, en opinión de Haidt, no es un razonamiento genuinamente abierto a las pruebas y razones, sino uno «motivado» por factores como el deseo de mantener buenas relaciones con los demás o evitar la ansiedad que puede producir ver amenazada nuestra propia visión del mundo. Por ello, el proceso de razonamiento es «más parecido a un abogado defendiendo a un cliente que a un juez o un científico buscando la verdad» (Haidt, 2012: 177), en el sentido de que no es un razonamiento imparcial: se buscan exclusivamente las pruebas a favor de nuestras creencias intuitivas («Una vez que las personas encuentran pruebas en su apoyo, incluso una parte mínima de una mala evidencia, a menudo suspenden la búsqueda, puesto que poseen una "epistemología de dar sentido" en la que el objetivo del pensamiento no es lograr la conclusión más atinada sino encontrar la primera conclusión que encaje con las creencias previas relevantes» (Haidt, 2012: 179).

3. Asimismo, y en tercer lugar, hay también experimentos neuropsicológicos (por ejemplo, los de Damasio) que pueden tomarse como indicios de que la conexión entre el razonamiento moral y la acción moral es bastante débil, y que las emociones (la empatía, fundamentalmente) resultan mucho más eficaces como motivadoras del comportamiento moral, dando la razón así a Hume, Adam Smith, etcétera.

De su intuicionismo y el papel residual del razonamiento, HAIDT obtiene tres conclusiones importantes:

a) La primera, que el proceso de razonamiento solo funciona objetivamente (y no meramente como una racionalización *a posteriori*) bajo circunstancias muy concretas: requiere que la persona tenga tiempo y capacidad, una motivación para ser exacto (orientarse hacia la verdad), que carezca de un juicio *a priori* que defender y que no estén presentes ni un interés por mantener la relación ni un interés en mantener la coherencia (HAIDT, 2012: 180). HAIDT no niega, por tanto, que en ocasiones el razonamiento pueda ser la causa de nuestros juicios morales («particularmente entre filósofos —escribe HAIDT—, uno de los pocos grupos que son capaces de razonar bien» (HAIDT, 2012: 172). Pero se trata de circunstancias poco frecuentes. En «circunstancias más realistas, el razonamiento moral no es libre para buscar la verdad» (HAIDT, 2012: 181). Y muchas veces, solo recurrimos al razonamiento con la intención de suscitar en el otro las mismas intuiciones que tenemos nosotros.

b) La segunda, que «nuestra vida moral está plagada de dos ilusiones»: la ilusión de que nuestro juicio moral se debe a nuestro razonamiento (la *ilusión de mover la cola del perro*) y la ilusión de que nuestros argumentos cambian las opiniones de los otros (la *ilusión de mover la cola del otro perro*: «es como pensar que forzar el movimiento de la cola de un perro, moviéndola con tu mano, hará feliz al perro») (HAIDT, 2012: 184).

c) Por último, la lección que extrae Haidt es que las raíces de la inteligencia humana no están en nuestra capacidad racional de buscar y evaluar pruebas, sino «en lo que la mente hace mejor: la percepción, intuición y otras operaciones mentales que son más rápidas, sin esfuerzo y generalmente bastante precisas» (HAIDT, 2012: 180).

Pero, ¿en qué consiste realmente una intuición? HAIDT —a diferencia de los filósofos intuicionistas que hemos visto antes—, trata de ser más preciso en la explicación de las intuiciones. A partir de los trabajos de diversos autores, propone las siguientes características distintivas entre el proceso de razonamiento y el proceso intuitivo (HAIDT, 2012: 170):

El proceso de la intuición Rápido y sin esfuerzo	El proceso de razonamiento Lento y con esfuerzo
No es intencionado y funciona automáticamente	Intencional y controlable
Es inaccesible; solo se tiene consciencia del resultado	Accesible a la consciencia y visible
No requiere prestar atención	Requiere atención, que siempre es limitada
Paralelamente distribuido	Secuencial
Coincidencia de patrones: el pensamiento es metafórico, holista	Manipulación simbólica: el pensamiento trata de preservar la verdad, es analítico
Común a todos los mamíferos	Propio de los humanos cumplidos los dos años de edad y de algunos simios entrenados de cierta manera
Depende del contexto	No depende del contexto
Depende de las bases (cerebrales y corporales) que lo sustentan	Independiente de sus bases (puede ser implementado en cualquier organismo o máquina que actúa conforme a ciertas reglas)

HAIDT se ocupa también de los mecanismos por los que se «fijan» las intuiciones y de su origen y desarrollo. Respecto de lo primero, echa mano de la hipótesis de DAMASIO de los marcadores somáticos y de la explicación de LAKOFF y JOHNSON del papel de las metáforas en nuestra manera de interpretar el mundo.

Como es sabido, de acuerdo con Antonio DAMASIO [DAMASIO, 2004: 201 y ss.] las experiencias del mundo se van asociando con sensaciones emocionales de placer o de dolor, de manera que quedan «etiquetadas» como positivas o negativas. Esto permite una rápida (y automática) evaluación de las situaciones análogas que se nos van planteando en el futuro. Llega un punto en el que «el mero pensamiento en una acción particular resulta suficiente para

provocar en el cerebro una respuesta de "como si", mediante el cual la persona experimenta de un modo débil los mismos sentimientos corporales que experimentaría si realizara la acción» (HAIDT, 2012: 190). Este «etiquetamiento» de situaciones como positivas o negativas le permite al cerebro descartar con gran rapidez las posibilidades de actuación asociadas a sensaciones negativas y seleccionar las asociadas a sensaciones positivas de una manera mucho más rápida que lo que puede hacer el análisis racional.

Por su parte, el lingüista George LAKOFF y el filósofo Mark JOHNSON han sostenido que «los procesos del pensamiento humano son en gran medida metafóricos» (LAKOFF y JOHNSON, 1995: 42). Las metáforas construyen un entramado de conceptos que determinan en gran parte nuestra manera de interpretar y vivir en el mundo. Por ejemplo, «una discusión es una guerra» («Tus afirmaciones son *indefendibles*», «*Atacó* todos los puntos débiles de mi argumento», «Sus críticas dieron justo *en el blanco*», «*Destruí* su argumento», «Nunca le he *vencido* en una discusión», «Si usas esa *estrategia*, te *aniquilará*») (LAKOFF y JOHNSON, 1995: 40); el tiempo es dinero («Me estás haciendo *perder* el tiempo», «Este artilugio te *ahorrará* horas». «No *tengo* tiempo para *dedicártelo*», «¿En qué *gastas* el tiempo estos días?», «He *invertido* mucho tiempo en ella», «¿Te *sobra* mucho tiempo?», «Vive de tiempo *prestado*») (LAKOFF y JOHNSON, 1995: 44); «lo racional es arriba» mientras que «lo emocional es abajo» («Mantuvimos una discusión de *alto nivel* intelectual»; «No pudo *sobreponerse* a sus emociones») (LAKOFF y JOHNSON, 1995: 54). Muchas de estas asociaciones metafóricas se construyen a partir de nuestras experiencias como cuerpos físicos: por ejemplo, asociamos la pureza y limpieza que necesitamos en los alimentos con la bondad moral, y la corrupción y suciedad con el vicio (HAIDT, 2012: 192). En definitiva, lo que nuestro autor parece querer sugerir es que estas asociaciones (que LAKOFF y JOHNSON parecen considerar automáticas y que, a su vez, determinan nuestros juicios) están detrás de muchas de nuestras intuiciones morales:

> La intuición moral, así, parece ser el producto automático de un conjunto subyacente y en gran medida inconsciente de conceptos morales interconectados. Estos conceptos pueden tener alguna base innata [...] la cual es entonces en gran parte construida por extensiones metafóricas desde la experiencia física. Las metáforas tienen

> vinculaciones y mucha de la argumentación moral y la persuasión implica tratar de conseguir que la otra persona aplique la metáfora correcta. Si Saddam Hussein es Hitler, esto implica que debe ser frenado. Pero si Irak es Vietnam, se sigue que Estados Unidos no debería implicarse en el conflicto. Tales argumentos son de hecho una forma de razonamiento, pero son razones diseñadas para provocar intuiciones en el oyente (HAIDT, 2012: 192).

Además del mecanismo de fijación de las intuiciones, HAIDT propone una hipótesis acerca del origen evolutivo de muchas de nuestras intuiciones morales. Su propuesta trata de explicarlas, como DARWIN, como un desarrollo de los instintos sociales de los animales. El proceso iría de las meras regularidades («normas descriptivas», las llama él) de muchas especies a las «normas prescriptivas» (normas cuyo cumplimiento viene reforzado por las reacciones de los otros) que aparecen en los primates y que con la aparición del lenguaje alcanzan un estadio superior [aunque «el sistema cognitivo y emocional de creación y refuerzo de normas estaba disponible mucho antes de que el lenguaje existiera» (HAIDT, 2012: 193)]. Se trata de normas relacionadas con la reciprocidad, la empatía, el altruismo, la lealtad, etc., y dado el paralelismo «entre la vida social de los humanos y la de los chimpancés, la carga de la prueba debe recaer sobre aquellos que quieren abogar por la discontinuidad, esto es, que la moralidad humana surgió *ex nihilo*, cuando desarrollamos la habilidad de hablar y razonar» (HAIDT, 2012: 194). Ahora bien, sobre esas intuiciones innatas que la evolución ha inscrito en los seres humanos, la sociedad ejerce una labor de «poda», modelándolas a través de la influencia de los demás, la inmersión en costumbres y los procesos de socialización (esto explica que los niños acaben con una moral propia de su grupo o comunidad). Sorprendentemente, HAIDT afirma que «una cultura que enfatizara todas las intuiciones morales que la mente humana está preparada para experimentar correría el riesgo de parálisis pues cada acción provocaría múltiples intuiciones en conflicto» (HAIDT, 2012: 196) (sorprendentemente, porque entonces la evolución no parece tan sabia ni tan adecuada para asegurar por sí sola la supervivencia por medio de pautas morales).

HAIDT propone incluso seis intuiciones innatas (y sus opuestas) como ejes de una «ética intuitiva» (lo que explicaría la coinciden-

cia de normas morales en diversas culturas): 1) Cuidado (daño); 2) justicia (engaño); 3) lealtad (traición); 4) autoridad subversión); 5) santidad (degradación), y 6) libertad (opresión) (*www.moral-foundations.org*). Combinaciones diferentes en el peso asignado a cada una de estas intuiciones o emociones darían lugar a los distintos sistemas morales (Haidt, 2004).

En definitiva, Haidt es, como vemos, uno de los autores que más ha reivindicado el papel de las intuiciones y las emociones en la formación de nuestras creencias y la toma de nuestras decisiones morales. Quizá, también, es uno de los autores que ha llevado su tesis más lejos, al entender que el razonamiento juega un papel muy reducido en la «cognición moral», salvo como racionalización *a posteriori* de nuestras decisiones. Su teoría plantea varios problemas (que veremos más adelante). Uno de ellos es, precisamente, la escasa relevancia que concede a la razón. ¿Es así en todos los casos? Parece más fácil aceptar que en algunos casos (como el del incesto, el aborto o la eutanasia), las emociones tienen un peso muy relevante, pero ¿ocurre lo mismo con todos nuestros juicios morales? Por otra parte, si los filósofos (como admite Haidt) pueden llegar a razonar de manera objetiva (orientada a la corrección) sobre dilemas morales, ¿es entonces una cuestión de educación y entrenamiento? ¿Qué diferencia hay entre los casos en los que el juicio moral depende de las intuiciones y emociones y los casos en los que la razón tiene una verdadera función decisiva? Joshua Greene ha propuesto, con ayuda de técnicas neurocientíficas, una respuesta a esta última pregunta.

2. J. GREENE Y LOS DILEMAS MORALES «PERSONALES»

El psicólogo y neurocientífico Joshua Greene es otro de los autores que ha estudiado el papel de las intuiciones en la toma de decisiones sobre cuestiones morales, aunque, a diferencia de Haidt, concede mayor peso al razonamiento en algunos tipos de juicios. Su objetivo es doble: en primer lugar, entender en qué casos los juicios morales dependen de intuiciones y en qué casos dependen del razonamiento; en segundo lugar, evaluar qué relevancia tienen las investigaciones de este tipo para la ética prescriptiva y la metaética.

El punto de partida de su hipótesis son los experimentos en neurociencia realizados por él mismo y sus colaboradores, descritos en el artículo An *fMRI Invesigation of Emotional Engagement in Moral Judgment* (GREENE *et al.*, 2001). GREENE y su equipo sometieron a un grupo de personas a dos conocidos dilemas morales [los «problemas de tranvías», propuestos inicialmente por Philippa FOOT (1994), que han llegado a derivar en una manera de hacer filosofía moral centrada en este tipo de supuestos artificiales], mientras observan su actividad cerebral a partir de técnicas de neuroimagen. Los dilemas fueron los siguientes:

1) Un tranvía, fuera de control, está a punto de atropellar a cinco personas. La única forma de salvarlos es presionar un botón que lo desviará hacia otra vía, pero entonces atropellará a una persona (en lugar de a los cinco).

2) Un tranvía, fuera de control, está a punto de atropellar a cinco personas. La única forma de evitarlo es empujar a un peatón obeso sobre la vía, matando a una persona pero salvando a cinco.

En experimentos anteriores se había encontrado que, mientras la mayoría de personas resuelven el primer dilema accionando el botón (pensando que está justificado o, incluso, que es un deber sacrificar a uno para salvar a cinco), en el segundo dilema la mayoría considera que empujar al hombre obeso estaría mal. Esto se había interpretado como indicio de una diferencia moral relevante entre los dos casos. Hasta aquí, nada nuevo. Lo que GREENE comprueba es que la actividad cerebral de los participantes en el experimento es muy distinta en uno y otro caso. Cuando los sujetos se enfrentan al segundo dilema, a diferencia de lo que ocurre cuando se enfrentan al primero, las zonas del cerebro relacionadas con las emociones se activan de una manera mucho más intensa. La hipótesis de GREENE es que esto se debe a que el primer dilema es más «impersonal» y el segundo más «personal»: las emociones incidirían más en los dilemas «cercanos y personales» (GREENE, 2012: 151 y 152). Esta hipótesis explicaría otros casos. Por ejemplo, mientras que la mayor parte de la gente consideraría inmoral no ayudar a un excursionista accidentado para evitar manchar de sangre la tapicería del coche, no consideraría incorrecto no responder a una carta de una ONG que pide una contribución para enviar ayuda médica a países donde es necesaria (GREENE,

2012: 152). «El dilema del excursionista herido —escribe GREENE— es un dilema "personal" en el que la violación moral se produce de manera "próxima y cercana". El dilema de la donación es un dilema "impersonal", en el que la violación moral no tiene ese carácter. Resumiendo una historia larga, encontramos que los juicios que respondían a dilemas morales "personales", comparados con los "impersonales", conllevaban una gran actividad en áreas cerebrales asociadas con la emoción y la cognición social» (GREENE, 2012: 153).

Dada la diferente manera de «trabajar» que tiene el cerebro en los dilemas personales y en los impersonales, GREENE concluye asumiendo la tesis del procesamiento dual (*Does moral action depend on reasoning? Less than it should*, en *https://www.templeton.org*). Nuestro autor usa una metáfora clarificadora: algunas cámaras de fotos tienen un mecanismo que permite hacer fotos automáticamente y otro que permite hacerlas manualmente. El mecanismo automático es más eficiente, pero menos flexible; el manual es más flexible, pero menos eficiente. De manera similar, el cerebro tiene un «diseño híbrido»: puede adoptar decisiones sociales de manera «automática» (es el mecanismo emocional) o de manera consciente y dirigida (es el mecanismo de la razón). El primer sistema puede llevarnos a error, pero es necesario para sobrevivir, porque el segundo sistema, aunque sea necesario para resolver problemas nuevos o especialmente complejos, puede ser en ocasiones demasiado lento. La sugerencia de GREENE es que el sistema de procesamiento emocional se activa en los dilemas morales personales. ¿Por qué? Su explicación tiene que ver —como en el caso de HAIDT— con nuestra historia evolutiva. Múltiples trabajos apuntan a que el altruismo entre los seres humanos ha surgido gracias a su utilidad evolutiva, pero entonces refleja el entorno de nuestros antecesores, no el nuestro: «Nuestros antecesores no evolucionaron en un entorno en el que gentes completamente extrañas en sitios opuestos del mundo podían salvar su vida haciendo sacrificios materiales considerablemente pequeños» (GREENE, 2012: 153), de manera que nuestros impulsos nos llevan a ayudar a los demás solo en situaciones cercanas y personales.

¿Qué relevancia moral tiene todo esto? GREENE acepta la separación entre «ser» y «deber ser»; aun así, señala que esto no priva

a la neuroética de conclusiones relevantes. Distingue entre su relevancia para la ética normativa y para la metaética. En el primer nivel, señala que debemos resistir la tentación de pensar que realmente hay una diferencia moral relevante entre los dos tipos de casos. Dada la imposibilidad de pasar del ser al deber ser, el hecho de que tengamos intuiciones distintas en los dilemas personales y los impersonales no quiere decir que las respuestas correctas para cada caso hayan de ser distintas. Es decir, no tiene por qué haber una buena razón detrás de cada intuición. El conocimiento que puede aportar la neurociencia aplicada al análisis del funcionamiento del cerebro ante los dilemas morales es, precisamente, mostrar que lo que hay detrás de muchas intuiciones no es un rasgo moral relevante desde el punto de vista de hoy en día, sino una manera de funcionar del cerebro forjada en contextos distintos de los actuales. La neurociencia puede ayudarnos a comprendernos mejor y, de esta manera, poner en duda algunas de nuestras intuiciones (veremos que Peter SINGER defiende una conclusión similar). De manera semejante, en el nivel de la metaética comprender el funcionamiento del cerebro puede contribuir a cambiar nuestras creencias acerca de la estructura de la moralidad. GREENE, por ejemplo, se pregunta si puede considerarse que el realismo moral es correcto, esto es, si existen hechos morales «ahí fuera» que nuestras intuiciones nos permiten captar: «Entender cómo elaboramos los juicios morales podría ayudarnos a determinar si nuestros juicios son percepciones de verdades externas o proyecciones de actitudes internas» (GREENE, 2012: 155). Su hipótesis es que «creemos en el realismo moral porque la experiencia moral posee una fenomenología perceptiva, y la experiencia moral tiene una fenomenología perceptiva porque la selección natural nos ha equipado con mecanismos para hacer juicios morales intuitivos basados en la emoción [...]. Podemos comprender nuestra inclinación hacia el realismo moral no como una penetración en la naturaleza de la verdad moral, sino como un producto de los procesos cognitivos eficientes que usamos para tomar decisiones morales» (GREENE, 2012: 156). ¿Es esto todo lo que podemos extraer de nuestro creciente conocimiento acerca del funcionamiento del cerebro cuando nos enfrentamos a problemas morales?

3. M. HAUSER Y LA *GRAMÁTICA MORAL UNIVERSAL*

Como hemos visto, GREENE es escéptico a propósito del realismo moral y cree que conocer cómo surgen nuestras intuiciones morales arroja serias dudas sobre la existencia de hechos morales objetivos. Pero puede sostenerse cierto grado de universalismo moral sin necesidad de postular hechos morales en un sentido tan fuerte. Sería posible sostener que nuestros cerebros están estructurados de manera que necesariamente compartamos algún núcleo común abstracto de valores, principios, creencias o ideas morales. En definitiva, ¿es posible fundamentar algún tipo de «universales morales» a partir de las investigaciones sobre psicología, biología y neurociencia aplicadas a la ética? Esto es lo que parece querer sugerir Marc D. HAUSER en su libro *La mente moral. Cómo la naturaleza ha desarrollado nuestro sentido del bien y del mal* (HAUSER, 2008). Su punto de partida es que «hemos desarrollado un instinto moral, una capacidad que surge de manera natural dentro de cada niño, diseñada para generar juicios inmediatos sobre lo que está moralmente bien o mal sobre la base de una gramática inconsciente de la acción. Una parte de esta maquinaria fue diseñada por la mano ciega de la selección darwiniana millones de años antes de que apareciera nuestra especie; otras partes se añadieron o perfeccionaron a lo largo de nuestra historia evolutiva y son exclusivas de los humanos y de nuestra psicología social» (HAUSER, 2008: 17). Hasta aquí, HAUSER se inscribe dentro de lo que podríamos llamar el «intuicionismo neurocientífico» (asumiendo cierto grado de «innatismo» moral combinado con influencias sociales), pero hay tres diferencias importantes frente a autores como HAIDT y GREENE. La primera, que esa capacidad innata, fuente de las intuiciones morales, no tiene en HAUSER una vinculación tan estrecha con las emociones (aunque tampoco con la razón, en el sentido habitual); la segunda, un debilitamiento consciente de la distinción entre «ser» y «deber ser»; la tercera, la creencia de que es posible referirse a una *gramática moral universal*, entendida como un conjunto de normas o principios básicos abstractos a partir de los cuales generamos nuestros juicios morales sobre los casos concretos. Centrémonos en ellas:

a) *Intuiciones, emociones y razón:* como hemos visto, tanto Haidt como Greene relacionan estrechamente las intuiciones con las emociones. Para Haidt las intuiciones morales se pueden explicar, siguiendo la hipótesis de Damasio de las emociones como marcadores somáticos que facilitan la respuesta ante determinados problemas, como atajos en la toma de decisiones morales; para Greene, las intuiciones surgen cuando el sistema de procesamiento emocional se activa ante los dilemas morales personales. Para Hauser, en cambio, la formación de las intuiciones es independiente tanto de la razón como de las emociones. Este autor distingue varios modelos de explicación del proceso de formación de los juicios morales, que pueden ser representados de la siguiente manera:

1) Hecho → Percepción del hecho → Razón → Juicio (Kant)
2) Hecho → Percepción del hecho → Emoción → Juicio (Hume)
3) Hecho → Percepción del hecho {Emoción / Razón} → Juicio (mixto)
4) Hecho → Análisis del hecho → Juicio {Emoción / Razón} (Rawls)

Como vemos, mientras para Kant y Hume son las razones o las emociones, respectivamente, las que generan el juicio moral, Hauser atribuye a Rawls la tesis de que el juicio moral es el resultado de un análisis intuitivo de la acción que juzgamos: «La percepción de una acción o un acontecimiento pone en marcha un análisis de las causas y consecuencias, la cual produce, a su vez, un juicio moral (lícito, obligatorio, prohibido). Las emociones, si es que desempeñan algún papel, lo hacen después del juicio [...]. Los análisis de la motivación o las intenciones que subyacen a una acción, junto con los análisis de las consecuencias buscadas y las previstas, proporcionan el material apropiado para nuestra facultad moral. La única función de las emociones es modular lo que efectivamente hacemos frente a lo que comprendemos o percibimos como moralmente permisible» (Hauser, 2008: 75). Este, el que atribuye a Rawls, es el modelo que a Hauser le parece que mejor describe la formación de nuestros juicios morales. A pesar de usar la expresión «análisis del

hecho» o «análisis de la acción», que parece sugerir una intervención activa y consciente de la razón, HAUSER sostiene que se trata de una operación instintiva. Lo que este autor sostiene es la existencia de un "sentido moral" o "instinto moral" independiente de las emociones que permite, a partir de una serie de principios básicos universales, desarrollar unos determinados contenidos morales. Ni el modelo de KANT (desarrollado en psicología por PIAGET y KOHLBERG, entre otros) ni el modelo de HUME (desarrollado por HAIDT y Frans DE WAAL, entre otros) logran explicar esta facultad, al dejar importantes interrogantes sin contestar[1].

b) *La falacia naturalista:* una segunda diferencia importante entre HAUSER, por un lado, y la línea de HAIDT y GREENE, por otro, es que estos dos últimos declaran aceptar la imposibilidad de pasar del ser al deber ser (otra cosa es que, en algunas de sus afirmaciones, no adviertan este deslizamiento) (HAIDT, 2012: 162; GREENE, 2012: 150). En el caso de HAUSER, por el contrario, explícitamente trata de argumentar que la falacia naturalista debe matizarse. En su opinión, aunque «los principios descriptivos que descubrimos en la naturaleza humana no tienen por qué guardar una relación causal (sic) con los principios prescriptivos [...] sin embargo, el corte quirúrgico entre hechos e ideales es excesivo» (HAUSER, 2008: 27). Su principal argumento parece ser el siguiente ejemplo, en el que «parece razonable partir del *es* para llegar al *debe*»:

> Hecho: la única diferencia entre un médico que aplica anestesia a una criatura y otro que no se lo aplica es que, sin ella, el niño sufrirá enormemente en el curso de una operación. La anestesia no tendrá ningún efecto pernicioso para el niño, sino que le hará perder temporalmente la conciencia y la sensación de dolor. Luego despertará, una vez acabada la operación, sin ninguna secuela negativa y con mejor salud, gracias al trabajo del médico.
>
> Juicio valorativo: por consiguiente, el médico debe administrar anestesia al niño (HAUSER, 2012: 28).

Como veremos más adelante, puede argumentarse que este ejemplo no constituye en realidad ninguna objeción a la tesis de la falacia naturalista ni de la guillotina de Hume. Se puede sostener

[1] Para la crítica de HAUSER a los dos modelos, véase HAUSER, 2008: 43 y ss.

que la conclusión del argumento construido por HAUSER o bien es un deber técnico, una necesidad práctica (y no genuinamente prescriptivo, por tanto), que se deriva de la descripción de un hecho junto con la premisa implícita de que el médico no desea causar sufrimiento innecesario al niño, o bien es un deber genuinamente prescriptivo, pero, de nuevo, se deriva de la descripción del hecho junto con una premisa implícita, de carácter normativo o valorativo, de acuerdo con la cual no se debe causar sufrimiento innecesariamente o es malo causar sufrimiento innecesariamente.

En todo caso, el intento de HAUSER muestra que su propósito no es meramente describir cómo funciona la facultad moral o el órgano moral del que afirma que estamos dotados, sino también prestar cierto apoyo justificatorio a los principios que constituyen la «gramática moral universal».

c) *Los elementos universales de la moral: la facultad o instinto moral.* Como hemos visto, la tesis central de HAUSER es que poseemos una facultad innata [en realidad, HAUSER habla de manera imprecisa de «instinto moral», «sentido moral», «capacidad moral» o, incluso, de «órgano moral» sin mayor aclaración (CORTINA, 2012: 242)] producto de la evolución «que de manera inconsciente y automática genera juicios sobre lo que está bien y está mal» (HAUSER, 2012: 26). Reconocer esta capacidad nos permite «entender mejor por qué algunos de nuestros comportamientos y decisiones se considerarán siempre injustos, permisibles o punibles» (HAUSER, 2012: 26).

HAUSER toma de RAWLS una analogía (que también puede encontrarse en Hare (HARE, 1999: 35)] entre las herramientas lingüísticas innatas (un conjunto de reglas) que, de acuerdo con CHOMSKY, nos permiten adquirir y organizar un determinado lenguaje y un conjunto de principios morales que establecerían el marco dentro del cual se desarrollarían los diferentes sistemas de moral social. RAWLS se refiere a esta cuestión después de proponer que debe superarse el intuicionismo moral tratando de dar una teoría que describa y dé coherencia a nuestras intuiciones. En palabras suyas:

> Supongamos que cada persona, después de cierta edad y dotada de la capacidad intelectual indispensable, desarrolla, en condiciones sociales normales, un sentido de la justicia. Adquirimos una habili-

> dad para juzgar las cosas como justas e injustas y para apoyar estos juicios en razones [...]. Ahora bien, podemos en principio pensar en la teoría moral (y subrayo la naturaleza provisional de este punto de vista) como el intento de describir nuestra capacidad moral; o, en el caso presente, podemos creer que una teoría de la justicia describe nuestro sentido de la justicia. Según tal descripción no se quiere simplemente ofrecer una lista de los juicios que estamos dispuestos a formular acerca de las instituciones y acciones, acompañados de las razones en que se apoyan en caso de que se ofrezcan. Más bien, lo que se requiere es una formulación de un conjunto de principios que, al ser conjugados con nuestras creencias y conocimientos de las circunstancias, nos condujesen a formular estos juicios junto con las razones en que se apoyan, si quisiéramos aplicar tales principios de manera consciente e inteligente [...]. En este punto resulta útil una comparación con el problema de describir el sentido de gramaticalidad que tenemos respecto a las oraciones de nuestra lengua materna. En este caso el objetivo es caracterizar la capacidad para reconocer oraciones bien formadas mediante la formulación de principios claramente expresados que lleven a cabo las mismas discriminaciones que el orador. Esta es empresa difícil que, aunque inconclusa, se sabe que exige construcciones teóricas que exceden en mucho a los preceptos *ad hoc* de nuestro conocimiento gramatical explícito. Es de suponer que una situación semejante se da en la filosofía moral (RAWLS, 2006: 55-56).

Aunque ciertamente RAWLS sugiere la analogía, hay que destacar que HAUSER va mucho más allá (hasta el punto de presentar un RAWLS distorsionado), pues RAWLS solo la expone de manera tentativa y, además, llama a los juicios que realizamos basándonos en esas intuiciones «juicios meditados» (alejándolos del carácter inconsciente y automático que les otorga HAUSER). En todo caso, HAUSER desgrana las características que, en su opinión, tiene esta «facultad moral» (HAUSER, 2008: 82 y 83):

1. La facultad moral consiste en un conjunto de principios morales que constituyen una «gramática moral universal».
2. Estos principios guían nuestros juicios morales, «pero no determinan de manera estricta nuestra manera de actuar».
3. Cada principio «genera un juicio rápido y automático» acerca del estatus deóntico de un acto o hecho.
4. Los principios son inconscientes.
5. Son innatos.

6. Constituyen el marco de los sistemas éticos particulares posibles.
7. Permiten adquirir el sistema moral de la propia cultura de una manera rápida y sin esfuerzo.
8. Los principios de la gramática moral universal son exclusivamente humanos.
9. La facultad moral debe estar en conexión con otras capacidades mentales (como el lenguaje o la memoria).
10. Las lesiones en las zonas cerebrales correspondientes a la capacidad moral pueden afectarla (las lesiones correspondientes a zonas auxiliares, como el lenguaje o la memoria, pueden afectar a la acción moral, en tanto que distinto a lo que piensan que deben hacer).

¿Qué pruebas tenemos a favor de la existencia de esta capacidad y estos principios morales innatos? HAUSER recurre a cuestionarios para sustentar sus afirmaciones (aún puede responderse a algunos de ellos a través de la web *http://www.moralsensetest.com/*). En estos cuestionarios plantea «dilemas tranviarios» y semejantes [«no solo tranvías, sino también elefantes en estampida, casas en llamas, botes de salvamento, administración de algunos medicamentos» (HAUSER, 2008: 163)] y pide a los encuestados que respondan qué harían ellos y justifiquen sus decisiones. De las respuestas infiere 1) que la mayoría de sujetos (un 90 por 100) responde que es lícito empujar la palanca para salvar a cinco personas, a costa de sacrificar a una, y solo un 10 por 100 responde que es lícito empujar a un hombre obeso, sacrificándolo para salvar a otras cinco personas; y 2) que prácticamente nadie puede justificar «suficientemente» por qué responde de manera distinta a uno y otro caso: «Entre las respuestas insuficientes —escribe HAUSER— había apelaciones a Dios, las emociones, las corazonadas, los sentimientos viscerales, las normas deontológicas (matar está mal), las consecuencias utilitarias (maximizar el mayor bien) y mi favorita: "¡por narices!". De entre los que juzgamos suficientes, la mitad hacían mención de algún aspecto del principio de doble efecto (no es lícito utilizar intencionalmente al hombre obeso para salvar a los cinco), mientras que otros se fijaban en la distinción entre daño personal e impersonal. El hecho de que la mayoría de la gente no tenga ni idea de por qué hace una distinción entre estos casos refuerza la tesis de que tienden a hacer juicios morales sin ser conscientes de los principios

subyacentes. Por otro lado, quizá los que invocan una justificación adecuada son mayores y más inteligentes, por lo general con buen nivel educativo y con formación específica en Filosofía moral y Derecho. Nada en nuestra muestra de sujetos apuntaba a semejante división» (HAUSER, 2008: 163 y 164).

Cabría esperar que el siguiente paso de HAUSER fuera el de señalar al menos algunos ejemplos de estos principios universales, pero no lo hace de una forma clara (¿quizá porque son inconscientes?). A veces parece estar pensando en criterios de reproche o valoración como el del doble efecto, la relevancia moral de las intenciones y las consecuencias, el peso de la mayor o menor inmediatez entre la acción y el resultado en el reproche Otras veces, en ideas abstractas como la equidad, la reciprocidad y la regla de oro de la moral; en otra ocasión propone como principio universal la obligación de cuidar a los niños, que es respetado en todas las sociedades (incluso aunque algunas de ellas admitan el infanticidio en determinadas circunstancias). Lo que considera universal son los principios, luego cada sociedad establece sus propias excepciones (HAUSER, 2008: 44).

En realidad, pese a que el trabajo de HAUSER pueda impresionar por su extensión y por la cantidad de datos aportados y pueda ser sugerente, muchas de sus tesis son difusas y, a veces, incluso abiertamente contradictorias. No solo no quedan claros los principios, sino tampoco el alcance de esta gramática universal: por un lado, presenta a los principios como meramente formales, vacíos de contenido [«estos principios *carecen de contenido específico*. No hay principios que dicten qué actos sexuales, altruistas o violentos son permisibles» (HAUSER, 2008: 486)], con lo que parece que es necesaria la concreción en una determinada moral social para ser efectivos; pero, por otro lado, una de las características de los principios de la gramática universal era que «cada principio genera un juicio *rápido y automático* acerca de si un acto o acontecimiento es moralmente lícito, obligatorio o prohibido» (HAUSER, 2008: 83). Tampoco queda claro si realmente su propuesta es meramente descriptiva [«este libro *describe* la manera en que nuestras intuiciones morales actúan y por qué han surgido» (HAUSER, 2008: 26)] o si tiene una pretensión prescriptiva o justificatoria de algunas creencias morales (como muestra su argumento a favor de la legitimidad

de pasar de hechos a valores en algunos casos). En fin, el libro de HAUSER podría verse como el primer paso de un proyecto, pero es bastante probable que este proyecto se vea truncado por las acusaciones de fraude científico que Hauser ha recibido y que, al parecer, han forzado su decisión de abandonar la investigación[2].

4. HACIA LA JUSTIFICACIÓN DE TEORÍAS, VALORES Y NORMAS MORALES A PARTIR DEL FUNCIONAMIENTO DEL CEREBRO. W. CASEBEER, PATRICIA CHURCHLAND Y M. S. GAZZANIGA

La propuesta de HAUSER de la existencia de una gramática moral universal no es la única que concede —si es que realmente lo hace— una dimensión normativa a las «implicaciones morales» de la neurociencia y la psicología. Otros autores lo hacen de una manera mucho más clara y abierta.

William D. CASEBEER, por ejemplo, cree que es posible mostrar que la neurociencia puede mediar, ni más ni menos, que en la discusión acerca de qué teoría de ética prescriptiva es más correcta: si el deontologismo de KANT, el consecuencialismo de John Stuart MILL o la ética de la virtud de ARISTÓTELES. El punto de partida de CASEBEER es que cada una de estas teorías contienen implícitas una psicología moral específica que exige capacidades cognitivas diferentes. Así, la teoría de KANT parecería requerir «al menos la capacidad de comprobar la consistencia lógica de máximas universalizadas de una manera independiente de la contaminación del afecto y la emoción» (CASEBEER, 2003: 842), capacidad que se correspondería con las funciones de la región frontal del cerebro. La teoría utilitarista de MILL requiere la capacidad de realizar cálculos utilitarios y el cultivo de emociones que nos muevan a

---

[2] La Universidad de Harvard, tras una investigación de casi tres años, le ha considerado culpable de hasta ocho casos de mala praxis científica (incluyendo falsificación o tergiversación de datos); posteriormente, ha habido también acusaciones de plagio en La mente moral. La Universidad de Harvard acordó suspender a HAUSER de sus clases en el 2011 y meses después este dimitió de su puesto. Es fácil encontrar tanto muestras de apoyo a HAUSER frente a la decisión de la universidad (considerándola infundada) como críticas al mismo.

procurar la felicidad de los demás (CASEBEER, 2003: 841), lo que implicaría las regiones prefrontal, límbica y sensorial del cerebro. La ética aristotélica de la virtud, por último, sería la más exigente, porque requiere educar nuestro carácter de manera que nuestros apetitos se coordinen con las buenas razones; esto implica una «psicología global» que requiere una intervención coordinada de las regiones del cerebro anteriormente mencionadas (CASEBEER, 2003: 842). Pues bien, nuestro autor cree —y lo argumenta de manera técnica a partir de experimentos neurológicos— que hay pruebas para aceptar, tentativamente, que la «cognición moral» pone en marcha de manera coordinada diferentes sistemas y redes cerebrales relacionadas tanto con la cognición como con las emociones (es decir, las regiones pre-frontal, frontal, límbica y sensorial: lo que podría llamarse «la zona de la cognición moral»), lo que muestra que «existe una clara convergencia entre la neuroética contemporánea y la psicología moral aristotélica» (CASEBEER, 2003: 845) y que «la psicología moral requerida por la teoría de la virtud es la más plausible neurobiológicamente» (CASEBEER, 2008: 841; CASEBEER y CHURCHLAND, 2003: 189). Podría pensarse que de aquí no se deriva nada sobre cómo debe ser el razonamiento moral, sino cómo de hecho razonamos sobre cuestiones morales. Pero debe tenerse en cuenta que KANT, MILL y ARISTÓTELES no trataban de describir cómo son los juicios morales, sino cómo deben ser. CASEBEER es consciente de esto y, de hecho, se defiende de las objeciones obvias de incurrir en la falacia naturalista de MOORE sosteniendo que esta se basa en la distinción entre proposiciones analíticas y sintéticas que hoy en día es muy cuestionada, y se defiende de la «guillotina de Hume» sosteniendo que lo que prohíbe es inferir deductivamente el paso del ser al deber ser, pero no por medio de otras inferencias no deductivas (como la inferencia de la mejor explicación) (CASEBEER, 2003: 482).

También Patricia CHURCHLAND ha argumentado en contra de la falacia naturalista de MOORE y la ley de Hume (tal como se ha entendido habitualmente). Patricia CHURCHLAND ha defendido la teoría de la mente conocida como materialismo eliminativo, esto es, la idea de que los estados mentales son estados cerebrales (físicos, por tanto) y de que los conceptos relativos a esos estados mentales, como intención, deseo, creencia, emoción, etc., deberán ser completamen-

te reelaborados o eliminados cuando la neurociencia arroje más luz sobre el funcionamiento del cerebro. Pero los cambios en la filosofía que provocará la neurociencia no se limitan a la teoría de la mente; en general, CHURCHLAND defiende el desarrollo de una «neurofilosofía», una de cuyas ramas sería la neuroética. En su opinión, es posible dar cuenta de la moral humana a partir de la evolución y el desarrollo de ciertas funciones cerebrales. En el origen de la moral es fundamental la idea de «apego», que está íntimamente relacionada con la hormona oxitocina. Nuestros cerebros están preparados para «preocuparse» (CHURCHLAND, 2012: cap. 2) por el bienestar propio, lo que constituye ya un valor básico. El siguiente paso, que constituye la transición a los valores sociales, es el cuidado de los demás: en primer lugar, de las crías y, posteriormente, de otros miembros de la familia, vecinos, etc., en círculos cada vez más amplios. Se trata de un proceso de «maternalización» de los individuos de la especie en la que la hormona de la oxitocina desempeña un papel relevante. A esta «fuerza» (el apego y cuidado de otros) se sumaron otras dos: el miedo y la ansiedad por la separación y una mayor capacidad de aprendizaje relacionada con el dolor y el placer. «En definitiva —escribe la autora—, la idea es que el apego —refrendado por el dolor de la separación y el placer de la compañía y gestionado por complejos circuitos neuronales y sustancias neuroquímicas— constituye la plataforma de la moralidad» (CHURCHLAND, 2012: cap. 2). La conclusión es que «la evaluación se asienta en las emociones y las pasiones que son endémicas de la naturaleza humana, así como en los hábitos sociales adquiridos durante la infancia [...]. La razón no crea valores, sino que se configura en torno a ellos y los lleva hacia nuevas direcciones» (CHURCHLAND, 2012: cap. 2). Por cierto, para nuestra autora el núcleo de la moral no son las normas o reglas, sino los valores [aunque a la luz de la filosofía del derecho actual, la distinción que traza entre unas y otros es bastante burda: considera que las reglas o normas no admiten excepciones mientras que los valores sí, y de ahí su rechazo de las primeras (CHURCHLAND, 2012: cap. 7)].

Lo que más llama la atención del libro de Patricia CHURCHLAND no es esta explicación del origen de la moralidad, que recuerda a la de HUME, sino su «puesta a la defensiva» frente a la objeción de que no se puede tratar el tema de los valores morales y la ética humana a partir del conocimiento científico. De esta manera, está reconociendo que su

pretensión no es meramente descriptiva, sino que es también un intento de justificar los valores surgidos de la manera descrita. CHURCHLAND considera un inmenso error la separación tajante entre descripciones y prescripciones, o entre hechos y valores, porque cierra la puerta a una naturalización de la ética (CHURCHLAND, 2012: cap. 2): «muchos filósofos morales creyeron que del mismo modo que la ciencia no puede decirnos nada esencial sobre lo bueno o lo valioso, tampoco puede decirnos nada sobre cómo deberíamos vivir. Podría ser informativa respecto a lo que cierta tribu piensa sobre la bondad, pero si esa bondad es genuina es siempre una pregunta abierta» (CHURCHLAND, 2012: cap. 2). Para CHURCHLAND la supuesta falacia naturalista se basa en la asunción falsa de que si afirmamos que «A es B» («lo bueno es la felicidad»), «A» y «B» deben ser sinónimos (cosa que, de acuerdo con ella, en ciencia no ocurre: podemos afirmar que «la temperatura es energía cinética a nivel molecular» sin que «temperatura» y «energía cinética a nivel molecular» sean sinónimos). Al igual que la salud «es un ámbito en el que la ciencia puede enseñarnos, y ya lo ha hecho, gran parte de lo que deberíamos hacer», también en «el ámbito de la conducta social [...] podemos aprender mucho de la observación común y de la ciencia acerca de las condiciones que favorecen la armonía y la estabilidad social, así como la calidad de vida individual». Tampoco la ley de Hume proscribe toda derivación del «ser» al «debe» (solo las derivaciones «facilonas» y deductivas): «En un sentido mucho más amplio de la palabra inferir que el de "derivar" se puede inferir ("dilucidar") lo que uno debería hacer, basándose en el conocimiento, la percepción, las emociones y el entendimiento, y comparando todas estas consideraciones entre sí. Lo hacemos constantemente, tanto en el mundo físico como en el social [...] ¿Tengo un dolor de muelas espantoso? Entonces debería ir al dentista ¿El horno está ardiendo? Debería apagarlo con bicarbonato de soda» (CHURCHLAND, 2012: cap. 1). En definitiva: «Nuestros cerebros, así como los cerebros de los animales en general, se organizan para valorar la supervivencia y el bienestar; la supervivencia y el bienestar son valiosos. Nuestras percepciones están impregnadas de valor, y en este sentido actúan como valencias» (CHURCHLAND: cap. 2). Por supuesto, volveremos sobre este punto.

En la misma línea de reclamar una filosofía (incluso una «filosofía de la vida», que nos diga cómo debemos vivir) basada en el cerebro encontramos a M. S. GAZZANIGA, uno de los autores más representativos y

conocidos de lo que podríamos llamar el «núcleo duro» de la neuroética (la neuroética no como mero saber auxiliar, sino como un saber omnicomprensivo). ARISTÓTELES, SÓCRATES, HUME, LOCKE, DARWIN, etc. han propuesto explicaciones de la naturaleza humana que, en muchos casos, han usado como palanca para proponer cómo debemos vivir. Pero «la cruda realidad es que esas ideas fértiles, metafóricas y atractivas —ya sean filosóficas o religiosas— son meros relatos», dice GAZZANIGA (2015: 168). Lo que la neurociencia puede aportar es, precisamente, un conocimiento realmente científico de la naturaleza humana. Una vez que la hayamos comprendido, podemos preguntarnos lo siguiente: «las verdades morales que aparentemente rigen nuestra vida ¿son un sistema de reglas independiente de nosotros, reglas que aprendemos y cumplimos? ¿O tales reglas surgen del empleo, por parte del cerebro, de sistemas inherentes de empatía con el fin de predecir la conducta ajena y actuar en consecuencias?» (GAZZANIGA, 2015: 179). La conclusión de este autor es «que no debemos buscar una ética universal que abarque verdades absolutas, sino la ética universal que nace del hecho de ser humano, que es claramente contextual, sensible a la emoción y orientada al refuerzo de la supervivencia» (GAZZANIGA, 2015: 179). En España, Francisco MORA ha asumido tesis análogas y postula «una ética universal, a través de la cual se puedan alcanzar valores y normas morales asumidos y respetados por todos los seres humanos. [...] Los valores éticos tan diferentes, para grupos étnicos tan diferentes, pueden converger en reglas y normas establecidas por la neuroética, basadas en el funcionamiento del cerebro humano» (MORA, 2007: 159). Ahora bien, dado que nuestros instintos morales tienden a la empatía hacia el cercano (como una cuestión de supervivencia), pero no a hacer el bien a aquellos con los que no tenemos contacto, MORA parece proponer que la vida social se reorganice en comunidades más pequeñas, la reducción de las ciudades, la vida en el campo..., como dictados de la nueva ética (CORTINA, 2010: 139).

5. ALGUNAS CONCLUSIONES DE LA DISCUSIÓN EN LA NEUROÉTICA

La neuroética, como hemos visto, ha retomado la discusión filosófica entre racionalismo y emotivismo moral, decantándose abru-

madoramente del lado del emotivismo y el intuicionismo. La mayoría de los autores examinados se encuentra a la sombra de HUME, poniendo a la empatía, el apego, la reciprocidad, etc. como fuente de las creencias morales. Generalizando los trabajos de los autores examinados, podríamos distinguir las siguientes tesis de la neuroética teórica, esto es, de la *neurociencia de la ética*:

a) La tesis del intuicionismo moral: la moral (nuestras creencias morales) es en gran parte intuitiva, lo que deja a la razón en un lugar muy secundario.
b) La tesis del emotivismo moral: las intuiciones morales proceden de las emociones (con la excepción de HAUSER, entre los autores que hemos visto).
c) La tesis del innatismo moral: las intuiciones son innatas y posteriormente son socialmente moldeadas.
d) La tesis del evolucionismo moral: las intuiciones morales tienen una explicación evolutiva.
e) La tesis de la normatividad de la neuroética: la neuroética puede justificar determinados valores o principios morales (no sostenida por todos).

Como hemos visto, el método de la neuroética se desarrolla en alianza con la ciencia empírica, es decir, tratando de vincular ética y experimentación o conocimiento científico, lo que confiere a sus tesis apariencia de gran solidez y fiabilidad. Sin embargo, todo resultado de un experimento debe ser interpretado, si se desea derivar consecuencias de él; y en esa interpretación es inevitable que se superpongan argumentos, conceptos, distinciones, etc., que deben resistir el análisis filosófico. En lo que sigue trataré de mostrar *a)* que las conclusiones descriptivas de la neuroética no son todavía tan sólidas como parecen, esto es, que para aceptarlas la neuroética debe responder primero a cuestiones que quedan abiertas y *b)* que la neuroética no tiene, por sí sola, implicaciones normativas directas, aunque sí puede contribuir a establecerlas combinada con las premisas (normativas o valorativas) adecuadas.

CAPÍTULO III

UN BALANCE CRÍTICO (I): INTUICIONISMO Y DARWINISMO MORAL

1. LA TESIS DEL INTUICIONISMO MORAL. INTUICIONES Y RAZÓN

Una primera conclusión en la que están de acuerdo todos los autores que hemos examinado es que nuestro juicio moral está dominado, o fuertemente influido, o condicionado en gran medida por intuiciones rápidas, inconscientes, no intencionales, ajenas a la reflexión. Ante los dilemas morales, nuestra respuesta suele ser intuitiva, irreflexiva, al menos en un primer momento. Las teorías psicológicas que, basándose en KANT, entendían que la razón era la fuente de nuestras creencias morales se equivocan, dado que le conceden un peso que en realidad no tiene. Podemos hablar, en este sentido, de un nuevo intuicionismo moral, basado en la interpretación de pruebas experimentales: un «intuicionismo neuroético». La otra cara de la moneda de esta tesis es que *la deliberación y discusión sobre los problemas morales no tiene relevancia genuina* (o la tiene muy menor, o solo en los casos en los que no nos vemos afectados personalmente). En la medida en que nuestras respuestas a los problemas morales vienen determinadas por las intuiciones, la razón queda desplazada y, por tanto, la deliberación y la discusión moral pierde importancia. Es más, en algunos autores, como HAIDT, su papel es —en la mayoría de las ocasiones— o bien una

racionalización *ex post,* o bien solo pretende causar en el oponente la misma intuición.

Esta concepción, sin embargo, puede estar incurriendo en una exagerada desconexión entre las intuiciones y la razón, esto es, parece estar presuponiendo que una excluye a la otra. Su carácter automático (y otras características, como la no intencionalidad y el ser un proceso inconsciente) parece alejar a las intuiciones del razonamiento. Sin embargo, hay varias vías para dar cuenta de las intuiciones de una manera más vinculada con el mismo. Mario BUNGE ha puesto de manifiesto que la idea de «intuición» engloba muchos fenómenos distintos. Si nos limitamos al ámbito de la ciencia, por ejemplo, «intuición» puede servir para referirse a *modos de percepción* (como la identificación rápida de objetos o acontecimientos, la comprensión clara del significado de signos o la capacidad de interpretación de fórmulas o diagramas), *formas de imaginación* (como la capacidad de representación de figuras geométricas o modelos abstractos, la capacidad para forjar metáforas o ver analogías o la imaginación creadora o capacidad de proponer hipótesis o conceptos nuevos), *formas de inferencia* (inferencias rápidas, a saltos, incompletas), *capacidad de síntesis* (facilidad de adquirir una visión global sobre un asunto) o *capacidad de evaluación* (frónesis) de una situación y de elección de las mejores alternativas (BUNGE, 2013: cap. III.1). Algunas de estas formas de intuición no pueden verse como opuestas o excluyentes de la razón: La inferencia rápida o incompleta es un razonamiento embrionario o primitivo y la «aprehensión sinóptica —como señala BUNGE— no es un sustituto del análisis, sino un premio al análisis esmerado» (BUNGE, 2013: cap. III.1). Pero, además, los diferentes tipos de intuición —incluso aquellos que no pueden verse en sentido estricto como razonamientos, ni siquiera incompletos— vienen favorecidos especialmente por el ejercicio continuado del razonamiento, del análisis sobre un problema, de la experiencia en una actividad o de la dedicación al estudio de una disciplina. Esto es así incluso en el caso de la imaginación creadora, de la pura corazonada, que podría considerarse la más alejada de la razón: Friedrich August KEKULÉ narró que resolvió el problema de la estructura del benzeno (que tiene forma de anillo) cuando, habiéndose quedado dormido frente a la chimenea de su casa, soñó con una serpiente que se mordía la cola, pero esto

solo ocurrió tras doce años de reflexionar sobre el problema. Son los expertos en un ámbito los que tienen más probabilidades de intuir hipótesis y soluciones acertadas. De acuerdo con estas consideraciones, la intuición no desplaza a la razón, sino que o bien es una forma imperfecta de ejercerla o bien esta es el campo más adecuado para que florezca. El pasatiempo de atribuir grandes descubrimientos científicos a intuiciones anecdóticas puede resultar entretenido, pero oculta el enorme esfuerzo de la razón que es necesario para que surja la chispa creadora.

También Peter GÄRDENFORS ha propuesto una explicación de las intuiciones que no las contrapone necesariamente a la razón. En su artículo *La intuición como conocimiento implícito* (2005) sostiene que las intuiciones son un tipo de conocimiento: un conocimiento implícito y especialmente difícil de explicitar, por no estar estructurado. Cuenta, como ejemplo, la extrañeza que le causaba su capacidad de intuir que su gata volvía a casa, a pesar de que no la veía ni oía, momentos antes de que esta apareciera por la ventana, hasta que se dio cuenta de que «había aprendido implícitamente que los pájaros daban llamadas de advertencia al aproximarse la gata, es decir, era el cambio en el canto de los pájaros lo que me llevaba a mirar por la ventana en busca de la gata» (GÄRDENFORS, 2005: 51). Para GÄRDENFORS las intuiciones son, entonces, un tipo de pensamiento o razonamiento no verbal (para él no todo pensamiento es proposicional) cuyo resultado experimentamos como una «percepción directa». Ahora bien, estas percepciones son dependientes de experiencias previas y de nuestro aprendizaje acerca de cómo categorizar el mundo. «Lo que experimenta un experto como percepción directa —escribe— depende de un largo entrenamiento. Los expertos amalgaman información en unidades significativas más amplias, expulsan información irrelevante y se acogen a un conjunto extenso de modelos mentales. Este procesamiento de información sensorial por el experto se hace implícitamente (inconscientemente), y para el experto parece "percepción directa de la verdad, hecho, etc. independiente de procesos de razonamiento", es decir, intuición» (GÄRDENFORS, 2005: 53). En definitiva, de acuerdo con este argumento la contraposición entre intuición y razón sería falsa; la intuición solo se contrapone al razonamiento explícito o proposicional. Las intuiciones a las que se refiere la neuroética podrían ser

de este tipo y la dificultad de dar razones que las justifiquen podría derivarse de la dificultad de acceder a ese conocimiento implícito y no proposicional, pero no de la ausencia de tales razones en la mente del sujeto.

Pero supongamos que aceptamos que los juicios morales proceden de corazonadas y que las corazonadas son independientes de los procesos de razonamiento. ¿Solo le quedaría a la razón la tarea de racionalizar *a posteriori* las decisiones tomadas, como piensa Haidt? En su famoso libro *Pensar rápido, pensar despacio* (Kahneman, 2012), Daniel Kahneman ha desarrollado —en sintonía con la neurociencia— la idea de que estamos dotados de dos sistemas de toma de decisiones, el sistema 1, rápido e inconsciente (el intuitivo) y el sistema 2, lento e intencional (el razonamiento). Se trata de una generalización de lo que los neuroéticos afirman para la moral. Kahneman, sin embargo, no relega al sistema 2 a un mero papel de autoengaño y racionalización. A partir de numerosos experimentos, sugiere que el sistema 1 suele ser suficiente en muchos casos, pero cuando «se ve en apuros» se activa el sistema 2: «La mayoría de las cosas que pensamos y hacemos (nuestro sistema 2) se originan en el sistema 1, pero el sistema 2 toma las riendas cuando estas cosas se ponen difíciles, y es él normalmente el que tiene la última palabra» (Kahneman, 2012, cap. 1). Es decir, si Kahneman tiene razón, la tarea de la razón no sería únicamente justificar la intuición *a posteriori*, sino también controlar la intuición cuando esta se produce. Además, tampoco habría que despreciar la razón «por llegar tarde»: el proceso de búsqueda de razones —aunque sea una vez ya tomada la decisión— tendrá seguramente influencia en *futuras* decisiones. Las razones se nos aparecerán como buenas o malas y ese proceso de reflexión dejará su huella. Las intuiciones pueden ser «huellas» de procesos de reflexión previos, procesos que, a su vez, pueden haber sido intentos de racionalizar decisiones tomadas intuitivamente.

En definitiva, parece haber razones para dudar de que las intuiciones desplacen todo papel causal de las razones en la toma de decisiones y realización de elecciones. La neuroética parece haber exagerado la separación entre intuición y razón y no haber advertido las relaciones entre ellas. Ello conduce a su vez a ver a las intuiciones o como ideas inflexibles o como percepciones determinadas por procesos causales no racionales, esto es, al determinismo en las

decisiones morales. Pero si aceptamos que la razón tiene un importante espacio para actuar, entonces podemos advertir la relevancia de la «educación en la razón»: si las intuiciones de los expertos son con más frecuencia acertadas, habrá que esforzarse en ser más expertos; si algunas intuiciones son «huellas» de razonamientos previos realizados para justificar *ex post* una intuición, habrá que esmerarse en encontrar la justificación de nuestras intuiciones (o descartarlas si no la tienen), aun cuando la decisión ya haya sido tomada; si son inferencias incompletas, habrá que ejercitarse en el razonamiento silogístico y profundizar nuestro conocimiento en el campo que se trate; si las intuiciones deben ceder el paso a la reflexión cuando advertimos determinadas «señales» de que el asunto es complejo, habrá que educarse en la percepción de esas señales y en incrementar el control del sistema 2 sobre el sistema 1, esto es, el control de las razones sobre la intuición.

2. LA TESIS EMOTIVISTA. INTUICIONES Y EMOCIONES

La mayoría —no todos— de los anteriores autores (y de otros muchos que no hemos tratado) vinculan las intuiciones con las emociones. El argumento implícito parece ser el siguiente: por una parte, las imágenes cerebrales demuestran que cuando razonamos moralmente se activan intensamente zonas del cerebro relacionadas con la emoción; por otra, los test psicológicos demuestran que en la mayoría de casos resolvemos los dilemas morales de manera intuitiva. Ambas cosas deben estar, por tanto, relacionadas. Las intuiciones surgen de las emociones, de la actividad de las zonas afectivas del cerebro. La tesis de DAMASIO de las emociones como «marcadores somáticos» que activan respuestas automáticas cuando nos encontramos ante situaciones análogas a las vivencias que generaron inicialmente esas emociones permite explicar la relación entre intuiciones y emociones. En este sentido, la mayoría de intuicionistas neuroéticos son también emotivistas (no todos: como hemos visto, para HAUSER las emociones no preceden al juicio moral).

De nuevo, la tesis de que las decisiones morales provienen de la emoción puede verse como un caso especial de la tesis general que enfatiza el papel de las emociones en la toma de decisiones (de

cualquier tipo). DAMASIO sugiere que el modelo clásico de toma de decisiones como el resultado de una ponderación de razones a favor o en contra de una u otra respuesta no es descriptivamente adecuado. Los seres humanos no deciden en realidad de esa manera (si así fuera, las decisiones racionales, simplemente, no serían posibles, porque no es posible procesar toda la información necesaria en el tiempo en el que disponemos para actuar). Las emociones nos ayudan a tomar decisiones racionales porque «filtran» las alternativas posibles en función de si en el pasado alternativas semejantes fueron «etiquetadas» como adecuadas (produjeron placer o satisfacción) o inadecuadas (produjeron frustración o dolor). Las emociones son, por tanto, «marcadores somáticos» de las posibles alternativas (DAMASIO, 2004: 201 y ss.).

Esta manera de reconstruir las decisiones (morales o de otro tipo), sin embargo, plantea, en mi opinión, varios problemas:

a) En primer lugar, la evidencia que tenemos del papel de las emociones en la toma de decisiones es, en realidad, indirecta. Lo que los neurocientíficos pueden comprobar es que cuando nos enfrentamos a un dilema moral las zonas del cerebro relacionadas con las emociones se activan de una manera especialmente acusada. Pero esto no permite inferir que las emociones generan el juicio moral (o la decisión que sea). Podría ser que el juicio moral generara la emoción: por ejemplo, darse cuenta, tras un análisis, de lo injusto de una situación puede generar indignación; o, en un dilema moral «tranviario», ser consciente de que cualquier solución causará un daño que puede causar pesar. De acuerdo con ARISTÓTELES las emociones se relacionan con las creencias en una doble dirección: por un lado, muchas emociones son generadas por creencias; por otro, muchas emociones generan o modifican nuestras creencias (GONZÁLEZ LAGIER, 2009: 26 y ss.). La neuroética parece estar viendo solo esta segunda conexión. Además, parece haber evidencia de que las emociones se relacionan más con la motivación de la acción moral que con el juicio moral (parece que los individuos psicópatas, a pesar de sus déficits emocionales, tienen intuiciones morales semejantes a los individuos «normales», pero no se ven motivados por ellas).

b) En segundo lugar, la noción de emoción como «marcadores somáticos» las convierte en «impulsos ciegos», meras sensacio-

nes que carecen de contenido proposicional (de intencionalidad), por lo que solo es posible dar una explicación causal de su relación con los juicios morales (y no una explicación teleológica o basada en razones). Esta manera de entender las emociones y su relación con la moral se aleja de las concepciones de la emoción más extendidas hoy en día entre los filósofos y muchos psicólogos. Para estas concepciones (que, en buena medida, reivindican la concepción de las emociones de ARISTÓTELES), debe distinguirse en ellas al menos tres dimensiones distintas: 1) Una dimensión cognitiva (en un sentido amplio, que incluye desde una creencia hasta una mera percepción); 2) una dimensión afectiva o puramente fenomenológica (la sensación de placer o dolor), y 3) una dimensión motivacional (una tendencia a la acción). Las teorías cognitivas de la emoción centran su atención en el primer aspecto, las teorías «somáticas» y mecanicistas de la emoción se centran en el segundo y las teorías conductistas en el tercer elemento. Por último, las teorías no reductivistas trata de dar cuenta de todos los aspectos de las emociones. Si se identifican las emociones con el segundo aspecto quedan muchos problemas por resolver: *a)* no se logra explicar la posibilidad de emociones inconscientes o sin sensación; *b)* no se logra explicar que las emociones puedan formar parte de explicaciones racionales (teleológicas), y no solo causales, de la conducta; *c)* no se logra explicar la relación entre las emociones y las creencias; *d)* no se logra dar cuenta de la posibilidad de evaluar a las emociones como razonables o no (en función de la creencia subyacente), y *e)* sobre todo, incurren —de nuevo— en un distanciamiento tajante entre las intuiciones de las que se dice que proceden nuestros juicios morales y la razón (GONZÁLEZ LAGIER, 2009, cap. II).

3. LA TESIS DEL INNATISMO MORAL

Autores como HAIDT, HAUSER, GAZZANIGA o CHURCHLAND apoyan el innatismo moral, esto es, la idea de que nacemos con valores, creencias o principios morales independientes en su origen del aprendizaje, de manera que la moral no es enteramente un producto cultural. Los argumentos a favor del innatismo descansan, de nuevo, en el carácter automático de la respuesta y en la incapacidad de

dar razones (HAIDT), junto con la coincidencia de respuestas a pesar de la heterogeneidad de los encuestados (HAUSER). Sin embargo, para asegurar que nuestras creencias morales son innatas hemos de descartar completamente que el automatismo se deba a la aceptación de principios profundamente arraigados, pero transmitidos culturalmente. Como señala Adela CORTINA comentando el ejemplo de HAIDT: «El hecho de que las personas encuestadas respondan de forma intuitiva, es decir, inmediata, automática, sin tener conciencia de cómo han llegado a formular el juicio, y que en muchas ocasiones no sepan dar razón de por qué una acción les parezca buena o mala puede muy bien explicarse porque lo han aprendido socialmente y no lo han sometido a revisión» (CORTINA, 2011: 86). Muchas normas de la moral social (como el incesto, en el ejemplo de HAIDT, si es que se acepta su carácter moral) son heredadas del entorno social, que las inculca la mayoría de las veces sin dar razones que las justifiquen. Por otra parte, la universalidad de las normas o principios también debe tomarse con precaución: muchos de los supuestos comportamientos morales universales no son, en realidad, de carácter moral (o exclusivamente de carácter moral): por ejemplo, como hemos visto HAUSER ha postulado la universalidad del principio de cuidado de las crías (que, obviamente, tiene una clara explicación evolutiva), pero ¿deberíamos extraer también la conclusión de que la búsqueda de alimento o el huir de los animales predadores tiene también valor moral? (BARTA, 2013: cap. III). ¿O la reproducción? Lo que quiero decir es que muchos hábitos aportados como ejemplos de conductas universalmente aceptadas no tienen (o no tienen exclusivamente) carácter moral.

En todo caso, lo que parece poner de manifiesto esta insistencia en el carácter innato o «instintivo» de las intuiciones morales es cierta confusión o falta de distinción entre habilidades o capacidades morales y contenidos (valores, principios, creencias) morales. Volveremos sobre esto en el siguiente epígrafe.

4. LA TESIS DEL EVOLUCIONISMO ÉTICO

Otra de las características del «intuicionismo neuroético» es su alianza con el darwinismo moral, que hace que la neuroética pue-

da verse como una continuación de la biología moral. Esta constituye otra diferencia importante entre el intuicionismo filosófico y el intuicionismo neuroético. Las intuiciones morales se ven como mecanismos que la evolución ha seleccionado porque aseguran la supervivencia de la especie. La conjunción de la tesis de la vinculación de las intuiciones con las emociones y la explicación evolutiva de las intuiciones morales podría hacer menos misteriosa la facultad de «intuir», que quedaba totalmente inexplicada en MOORE, ROSS, etc., pero, como veremos, se ha de enfrentar a serias dificultades.

Lo primero que hemos de preguntarnos es de qué se está hablando cuando se afirma que la moral es un producto de la evolución: ¿de la capacidad de valorar las acciones desde un punto de vista moral o de los principios, valores o reglas morales? Como ha señalado Francisco AYALA, «la cuestión de si el sentido ético está biológicamente determinado puede desdoblarse en los dos problemas siguientes: 1) ¿está determinada por la naturaleza biológica del ser humano la capacidad para la ética?, y 2) ¿están determinados biológicamente los sistemas o códigos de normas éticas aceptadas por los seres humanos?» (AYALA, 2013: 61). La respuesta afirmativa a la primera pregunta no condiciona la respuesta de la segunda. Es posible afirmar que el hombre es un ser ético (es decir, un agente que juzga las acciones como correctas o no desde un punto de vista normativo) por naturaleza y negar que existan normas éticas determinadas biológicamente (esta es la postura de AYALA): «La necesidad de tener pautas morales nada nos dice acerca de cuáles serán dichas pautas, lo mismo que la capacidad para el lenguaje no determina qué idioma hablaremos» (AYALA, 2013: 61).

De acuerdo con este autor, existen tres condiciones para el comportamiento ético: 1) La capacidad de anticipar las consecuencias de las propias acciones, 2) la capacidad de hacer juicios de valor y 3) la capacidad de escoger entre líneas de acción alternativas. La primera está vinculada con la habilidad para relacionar medios y fines, que surgió con la posibilidad de construir y usar utensilios a partir del momento en que se adoptó la postura erguida y las manos se convirtieron en órganos manipuladores. La segunda está vinculada evolutivamente con la capacidad de abstracción, que permite clasificar hechos y objetos en categorías generales, lo cual a su vez permite compararlos. La tercera es consecuencia del desarrollo

de una inteligencia que permite imaginar y explorar mentalmente distintas líneas de acción (AYALA, 2013: 62 y ss.). Ahora bien, es importante en el pensamiento de AYALA (frente a la hipótesis de la neuroética) la idea de que el comportamiento ético no es directamente un resultado de la evolución, sino solo indirectamente y en la medida en que es una consecuencia del desarrollo de la inteligencia humana; es decir, lo que tiene valor adaptativo y ha sido favorecido por la evolución es la inteligencia humana, no el hecho de ser capaz de comportarse moralmente (que es consecuencia a su vez de la inteligencia humana): «No encuentro argumentos a favor de que el comportamiento ético se desarrollara por ser intrínsecamente adaptativo. No veo cómo la valoración de ciertas acciones como buenas o malas (no la mera elección de unas y no otras, o la decisión respecto de sus consecuencias prácticas) incrementaría la eficacia reproductora del valorador. Ni tampoco veo cómo podría existir una forma de comportamiento ético incipiente que fuera luego fomentada por selección natural» (AYALA, 2013: 66).

A diferencia de la capacidad para la moral, los «códigos morales», de acuerdo con AYALA, no son determinados por la biología, sino que su carácter es cultural. Entre ellos y nuestra naturaleza biológica suele haber cierta coherencia, porque «los sistemas morales, como cualquier otra actividad cultural, no pueden sobrevivir mucho tiempo si discurren en franca contraposición con nuestra biología», pero esta coherencia no tiene por qué darse necesariamente (AYALA, 2013: 69). En su opinión, las mismas condiciones que hacen posible el comportamiento moral nos capacitan para no aceptar como moralmente correcto cualquier comportamiento biológicamente determinado. Pensar que hemos de aceptar como correcto todo comportamiento favorecido por la evolución es violar la ley de Hume (AYALA, 2013: 73): «La naturaleza biológica es moralmente neutra» y sin premisas valorativas o normativas no podemos derivar conclusiones normativas.

Los intentos de fundar *contenidos* morales en la selección natural buscan un tipo de comportamiento o actitud que se considera evolutivamente valioso (esto es, beneficioso para la supervivencia de la especie), como la empatía o el altruismo o la capacidad de reciprocidad, y tratan de derivar de él los principios o normas de lo que consideramos códigos morales aceptables. Esta estrategia no

es exclusiva de la neuroética, sino que es continuación de propuestas que se remontan (en la época moderna) hasta DARWIN y que han sido retomadas por autores como SPENCER, WILSON, HUXLEY, WADDINTON, etc. Tomemos, por ejemplo, la explicación del altruismo de Michel RUSE. Una conducta es altruista cuando se realiza para el beneficio de otro, aunque no suponga ningún beneficio —o incluso implique un sacrificio— para el sujeto que se comporta de esa manera. Podemos distinguir entre un altruismo biológico y un altruismo moral. El primero se realiza «porque los genes que favorecen tal comportamiento son realmente favorecidos por la selección natural [...], aunque la aptitud del individuo que manifiesta el comportamiento decrezca» (AYALA, 2013: 79) (es necesario tener en cuenta que para sopesar los efectos de la selección natural no basta con tener en cuenta las ventajas evolutivas de un rasgo para un individuo determinado, sino sobre todos los portadores del gen que causa ese rasgo). En relación con este altruismo, hay suficientes pruebas de que la cooperación entre los seres humanos en la búsqueda y obtención de alimentos fue esencial para la supervivencia de unos grupos o comunidades. El altruismo moral, por el contrario, tiene una motivación ética: no el beneficio genético, sino la consideración por los demás. Pues bien, lo que RUSE sostiene es que el «altruismo biológico» (la necesidad de cooperación) es tan esencial para los seres humanos que «la naturaleza nos ha llenado de ideas sobre la necesidad de cooperar» (RUSE, 2004: 671 y ss.): «Pensamos que debemos ayudar, que tenemos obligaciones para con los demás, porque tener estas ideas va en nuestro interés biológico. Pero desde una perspectiva evolutiva estas ideas existen sencillamente porque aquellos de nuestros antepasados que las tuvieron sobrevivieron y se reprodujeron mejor que los que no. En otras palabras, el altruismo es una adaptación humana, igual que lo son nuestras manos y ojos y dientes y brazos y pies. Somos morales porque nuestros genes, modelados por la selección natural, nos llenan de ideas sobre la conveniencia de serlo» (RUSE, 2004: 672). De manera abreviada: somos altruistas en sentido moral porque necesitamos ser altruistas en sentido biológico.

Uno de los problemas de anclar la moral en el altruismo es el hecho de que, como advirtió HUME, no parece existir en nosotros un sentimiento de benevolencia o altruismo —en sentido moral—

universal. El altruismo biológico, la cooperación y la reciprocidad evolutivamente útiles son los que se orientan a los que comparten los mismos genes, al grupo o a la comunidad cercana, pero no al extraño. Como hemos visto que señaló GREENE: «consideremos que nuestros antecesores no evolucionaron en un entorno en el que gentes completamente extrañas en sitios opuestos del mundo podían salvar su vida haciendo sacrificios materiales considerablemente pequeños. Tengamos en cuenta también que nuestros antecesores evolucionaron en un entorno en el que en ocasiones los individuos que tenían entre ellos una relación cara a cara podían salvar la vida unos de otros solo a través de un sacrificio personal considerable. Dicho esto, tiene sentido que hayamos desarrollado instintos altruistas que nos llevan a ayudar a otros en extrema necesidad, pero la mayor parte de las veces cuando los necesitados están presentes de un modo "cercano" y "personal"» (GREENE, 2012: 153). ¿Debemos renunciar a la idea de una moral universal o confiar en que estos sentimientos universales se extiendan? En palabras de Adela CORTINA: «Es preciso resolver la disonancia entre esos supuestos juicios intuitivos que nos llevan a apostar por los cercanos y la pretensión de una moral universal como la que defiende la Declaración de Derechos Humanos de 1948. Como veremos, la sugerencia de que se ha desarrollado ya o se desarrollará una supuesta benevolencia universal que nos hará cordialmente cosmopolitas es un sueño hermoso, pero sin bases empíricas. Cosa que no puede permitirse una ética que pretende basarse en evidencias científicas» (CORTINA, 2011: 127).

Otro problema del argumento evolutivo es que podemos encontrar tanto ejemplos de normas morales que no parecen tener valor adaptativo (como muchos derechos humanos) como comportamientos biológicamente determinados cuyo valor moral es ampliamente rechazado (se ha sugerido, por ejemplo, el valor adaptativo de la conducta territorial, de los instintos agresivos, de considerar más grave el adulterio de la esposa que el del esposo, etc.), lo que nos lleva al problema central de la neuroética y de cualquier ética natural: la normatividad.

CAPÍTULO IV

UN BALANCE CRÍTICO (II): EL PROBLEMA DE LA NORMATIVIDAD

1. EPISTEMOLOGÍA NATURALIZADA Y NEUROÉTICA. ALGUNOS PARALELISMOS. EL PROBLEMA DE LAS FICCIONES ÚTILES

El proyecto fuerte de naturalización de la filosofía, esto es, su reducción a las ciencias empíricas, se enfrenta con un reto importante cuando se trata de «naturalizar» ámbitos de la filosofía que tengan una dimensión normativa. Se trata de una limitación grave, porque podría pensarse que la normatividad y la justificación son pretensiones intrínsecas a amplios sectores de la filosofía, si no a toda ella: la lógica, por ejemplo, pretende señalarnos qué razonamientos son correctos. La estética, cuáles son los cánones de belleza. La epistemología, qué métodos de conocimiento están justificados. La ética, cómo debemos vivir. Podemos tratar de recluir la reflexión filosófica en un metalenguaje dedicado al análisis conceptual de las nociones implicadas en el nivel normativo (conceptos como los de «implicación», «armonía», «verdad» o «bondad»), pero las relaciones entre este nivel y el nivel normativo son tantas que solo ilusoriamente logramos eludir la normatividad, aunque este es un punto que no puedo entrar a discutir aquí.

Tomemos la propuesta de QUINE de «naturalizar» la epistemología, que ha seguido una ruta semejante al recorrido de la bioética y

la neuroética. Como señala Antonio DIÉGUEZ: «La epistemología o teoría del conocimiento tradicional presuponía que la tarea de analizar el conocimiento humano, en la medida en que era una tarea filosófica, podía hacerse de un modo *a priori*, ya fuera mediante el análisis conceptual, lógico, o trascendental; sin necesidad, por tanto, de recurrir a los resultados de la investigación empírica» (DIÉGUEZ, 2003: 1). Frente a ello, QUINE propuso que la epistemología debía integrarse en las ciencias empíricas, en particular la psicología, abandonando su método especulativo y analizando empíricamente los métodos efectivamente usados por las personas para justificar sus creencias. Una parte de los estudios dedicados a esta tarea ha desembocado en una «epistemología evolucionista», cuya tesis central describe PUTNAM de la siguiente manera: «La razón es una capacidad que tenemos para descubrir verdades. Dicha capacidad tiene valor de sobrevivencia; evolucionó de la misma forma en que cualquiera de nuestros órganos o capacidades físicas lo hicieron. Una creencia es racional si se llega a ella mediante el ejercicio de esta capacidad» (PUTNAM, 2007: 194). Desde una perspectiva evolucionista hay muchas preguntas que plantearse sobre la mente humana y sus capacidades: «¿Cuál es la historia evolutiva de la mente humana? ¿Qué papel jugaron en dicha evolución los factores puramente biológicos (aumento del tamaño del cerebro, cambios en la estructura neuronal) y los factores culturales (lenguaje, uso de herramientas)? ¿Esta evolución fue una respuesta adaptativa al ambiente natural o al ambiente social? ¿Desde un punto de vista evolutivo, es mejor concebir la mente como una estructura modular, en la que cada módulo ha evolucionado independientemente y para propósitos específicos, o como un programa de propósito general? En otras palabras, ¿se parece la mente más a una navaja suiza o a un ordenador? ¿Cuáles son las diferencias más importantes entre la mente humana y las mentes de otros primates superiores?» (DIÉGUEZ, 2003: 4). Ahora bien, junto a estas cuestiones empíricas, surgen dos importantes preguntas de marcado acento filosófico, una epistemológica (y normativa o justificatoria) y otra ontológica. La pregunta epistemológica es la siguiente: si nuestras capacidades cognitivas y nuestras maneras de conocer tienen valor adaptativo, ¿quiere ello decir que son fiables, esto es, están justificadas nuestras creencias basadas en ellas? La respuesta de QUINE era claramente afirmativa: «Las criaturas que yerran inveteradamente en sus in-

ducciones tienen una tendencia patética, pero encomiable, a morir antes de reproducir su clase» (QUINE, 1970: 48). Sin embargo, la afirmación de que nuestras capacidades cognitivas son útiles no implica necesariamente que la realidad se ajuste a nuestras creencias: creencias útiles no tienen por qué ser creencias verdaderas. No está claro que la verdad sea más adaptativa que la falsedad: pensar en el enemigo como un ser infrahumano motiva más en la lucha que la idea de que es igual a nosotros, por lo que, si dicha creencia viniera dada por una mutación genética, sus portadores dejarían más descendencia (DIÉGUEZ, 2002: 18). Puede ser, por tanto, que lo que tenga valor adaptativo sea la capacidad de formar esquemas con eficacia motivadora, o explicativa y predictiva, sin que estos se correspondan con la realidad. Pensemos en el siguiente ejemplo de Antonio DIÉGUEZ:

> El fogonero del viejo vapor en el que Marlow, el protagonista de la célebre narración de Conrad, remontaba el inmenso río, adentrándose cada vez más en la selva, era «un muchacho realmente eficaz». Sabía encender y hacer funcionar la caldera a la perfección. Y, sin embargo, sus conocimientos técnicos eran nulos. Se trataba de un salvaje bien entrenado que había forjado para sí mismo una curiosa explicación de su actividad: «lo que sabía era que, si el agua desaparecía de aquella cosa transparente, el mal espíritu encerrado en la caldera mostraría su cólera por la enormidad de su sed y tomaría una venganza terrible». Sin duda un experto en termodinámica habría explicado el asunto de una manera distinta, pero no habría hecho que el barco avanzase mejor entre el silencio de la espesa vegetación. La teoría del fogonero, una superstición ingenua, le servía para controlar la presión mediante un indicador al efecto, para relacionar causalmente el nivel del indicador con lo que sucedía en el interior de la caldera. Si no era capaz de entender los conceptos de evaporación, presión o trabajo, ni de formarse una imagen precisa de lo que ocurría, ¿a qué desengañarlo? A la hora de la verdad el resultado era tan bueno como si conociera los principios científicos en los que se basaba su funcionamiento, e incluso puede que el temor al espíritu prisionero le hiciera más diligente (DIÉGUEZ, 1998: 9).

La explicación del fogonero, aunque falsa, es tan útil para hacer avanzar la embarcación como la del científico. Por supuesto, podemos pensar que la explicación científica es mejor que la del fogonero, porque seguramente permitirá un campo de predicciones

más amplio (y en relación con un mayor abanico de fines), pero ¿cómo podemos saber que no se trata simplemente de otra ilusión más sofisticada, sin ninguna conexión con la realidad? ¿No podría ser que (muchas de) nuestras creencias fueran ficciones útiles, mas no verdaderas? Supongamos que un evolucionista contesta que podemos distinguir entre creencias verdaderas y meras ilusiones útiles aplicando los métodos que la evolución nos ha dado. Su respuesta sería, entonces, claramente circular, porque estaría fundamentando su confianza en los métodos determinados evolutivamente a partir de los métodos determinados evolutivamente. De manera que la aceptación de la utilidad evolutiva de nuestras creencias no nos libra del problema de buscar criterios normativos independientes para confirmar qué conocimientos están justificados (en el sentido de ser verdaderos, no de ser útiles) y qué métodos son correctos para aprehender la verdad.

La segunda pregunta —la pregunta ontológica— que la epistemología evolucionista plantea está estrechamente relacionada con la primera. ¿Permite la epistemología evolucionista afirmar la existencia de una realidad independiente? Si nuestras creencias tienen valor adaptativo es porque son verdaderas, y si son verdaderas es porque se corresponden con la realidad en la que intentamos sobrevivir. Luego, «la realidad es real». Es más, debe haber una realidad, independiente de nuestras creencias (no construida por ellas, o no por completo) a la que adaptarse para sobrevivir. La epistemología evolucionista apoyaría cierto realismo ontológico, aunque podría ser compatible con un buen grado de constructivismo (respecto de la clasificación de los objetos del mundo, por ejemplo). No se trata tanto de una prueba o demostración del realismo como de una presuposición razonable del mismo (y que hoy en día casi nadie negaría), pero en todo caso está también sujeta a la idea de que la evolución nos asegura la fiabilidad aproximada de la mayor parte de nuestras creencias, cosa que, como hemos visto, puede discutirse.

En definitiva, la epistemología naturalizada no puede escapar del problema de la normatividad. El hecho de que la evolución haya favorecido nuestras habilidades cognitivas no asegura que nuestras creencias sean verdaderas. Lo único que hace es explicar por qué las tenemos, no justificarlas.

Si trazamos el paralelismo entre la epistemología evolucionista y la ética evolucionista que subyace a la neuroética nos encontramos también con la pregunta ontológica y la pregunta epistemológica: ¿existen hechos morales independientes de nuestras creencias, intuiciones y emociones? (esto es, ¿avala la neuroética el realismo moral?). Y ¿son las intuiciones y emociones que hay tras nuestros juicios morales correctos (verdaderos)? (esto es, ¿es la neuroética una teoría cognoscitivista?). Como en el caso de la epistemología naturalizada, las respuestas no son evidentes. Podría pensarse —continuando con la analogía— que si la evolución ha favorecido ciertos criterios para solucionar dilemas morales es porque estos dan las soluciones correctas a esos problemas, y si ha favorecido la creencia en ciertos valores, principios o normas es porque son genuinas obligaciones. Por tanto, existen soluciones correctas, normas o deberes independientes de nuestras creencias (realismo moral) y podemos conocerlos (cognoscitivismo) (CASEBEER y CHURCHLAND, 2003: 188). El argumento resaltaría la importancia de la vida en sociedad para la supervivencia y la importancia de la moral para la vida en sociedad. Ya hemos visto algunos problemas de esta línea de razonamiento (la distinción entre capacidad para la moral y contenidos morales, la posibilidad de que la capacidad para la moral sea un derivado de la inteligencia, pero sin valor evolutivo directo, y la dificultad de pasar de reglas prudenciales acerca de lo mejor para sobrevivir a genuinas normas morales, esto es, el problema de fundar un altruismo universal). Pero, además, nos tropezamos de nuevo con la objeción que vimos a propósito de la epistemología: ¿cómo sabemos que la evolución no ha favorecido la ilusión de que la moral es objetiva, de que hay respuestas correctas (verdaderas) a los problemas morales y genuinos deberes éticos? ¿No podría ser, simplemente, que se trate de una ilusión útil? Esta es la postura de RUSE: «¿No sucede en ocasiones que cuando uno ha ofrecido una explicación causal de ciertas creencias puede ver que estas, en sí mismas, ni tienen fundamento ni podrían tenerlo? [...] Tan pronto percibimos que nuestras creencias morales no son más que una adaptación establecida por selección natural para fomentar nuestros fines reproductores, esto es el fin de la moralidad. La moralidad no es más que una ilusión colectiva que nos han endosado nuestros genes para fines reproductores» (RUSE, 2004: 675). O sea, que estamos

determinados a considerar nuestras creencias morales como obligatorias, lo cual no quiere decir que lo sean. Esta es también, como vimos, la postura de GREENE: «Creemos en el realismo moral porque la experiencia moral posee una fenomenología perceptiva, y la experiencia moral tiene una fenomenología perceptiva porque la selección natural nos ha equipado con mecanismos para hacer juicios morales intuitivos basados en la emoción [...]. Podemos comprender nuestra inclinación hacia el realismo moral no como una penetración en la naturaleza de la verdad moral, sino como un producto de los procesos cognitivos eficientes que usamos para tomar decisiones morales» (GREENE, 2012: 156).

2. PETER SINGER Y LAS INTUICIONES SOSPECHOSAS

Pero no es necesario socavar todo intento de fundamentación objetiva de la ética para discutir la normatividad de las intuiciones morales. La ética de Peter SINGER es, en varios sentidos, una ética naturalista: por un lado, en el sentido de que, como utilitarista de preferencias, sostiene que «bueno» o «debido» puede entenderse en términos de la propiedad natural «lo que satisfaga la mayor cantidad de preferencias del mayor número de seres sintientes». Por tanto, SINGER no es un escéptico moral. Por otro lado, también en el sentido de que acepta que la moral tiene, o puede tener, una base evolutiva (por tanto, no rechaza el interés de la neuroética): «Nuestra biología no prescribe las formas específicas que adopta nuestra moralidad. Hay variaciones culturales en la moralidad humana, como sabía hasta Heródoto. No obstante, parece probable que todas estas formas diferentes sean la consecuencia de la conducta que existe en los animales sociales, como resultado de los procesos evolutivos usuales de selección natural» [Singer, 2012:285]. Pero ello no le lleva a aceptar una ética evolucionista de contenidos. Se puede estar de acuerdo con la explicación de la moral que ofrece la neuroética o la bioética sin dar el salto a considerar que la evolución justifica nuestras intuiciones morales. Por el contrario, comprender cómo han surgido nuestras intuiciones morales (su carácter contextual, dependiente de situaciones en muchos casos ya superadas, y su relación con

reacciones emocionales) «sugiere que no hay razón para intentar encontrar principios morales que justifiquen las diferentes intuiciones a las que se llega en los casos dados» (SINGER, 2012: 297). Estos principios pueden no existir y, si existen, su fundamento no descansa en la evolución. Este punto es especialmente importante para Peter SINGER, porque un argumento frecuentemente esgrimido contra el utilitarismo es que en ocasiones resulta contrario a nuestras intuiciones morales y porque una de las concepciones de la justicia más influyentes hoy día (la de John RAWLS) presta una atención especial a las intuiciones, que son el contrapeso en la balanza del equilibrio reflexivo de nuestras teorías. Si no hay razón para conceder tal peso a nuestras intuiciones (y eso es lo que muestran, según SINGER, la bioética y la neuroética), la teoría de la justicia de RAWLS está aceptando un lastre que la convierte en conservadora e irracional: «Hay escasas razones para construir una teoría moral diseñada para igualar juicios morales tomados en consideración y que propiamente provienen de nuestras respuestas que han evolucionado a partir de situaciones en las que nosotros y nuestros antepasados vivieron durante el periodo de nuestra evolución como mamíferos sociales, primates y, finalmente, seres humanos» (SINGER, 2012: 298). Lo que SINGER propone finalmente no es, sin embargo, desechar todas nuestras intuiciones (admite que no es posible hacerlo), sino pasarlas por el tamiz de la razón para «separar aquellos juicios morales que debemos a nuestra historia evolutiva y cultural de aquellos que tienen un fundamento racional» (SINGER, 2012: 301). Ahora bien, si se acepta esto, la propuesta de SINGER acaba acercándose a la de RAWLS, porque, ¿no será el método para esta selección de intuiciones contrastarlas con las teorías? ¿Y no es esto lo que proponía RAWLS? La diferencia entre uno y otro acaba siendo de grado: SINGER parece más propenso a abandonar las intuiciones que no cuadran con la teoría, y RAWLS parece más propenso a modificar las teorías para hacerlas cuadrar con las intuiciones (o así lo interpreta SINGER).

En todo caso, lo que anda rondando toda la argumentación de SINGER es que no hay razones para aceptar el valor normativo de nuestras intuiciones morales solo porque hayan sido desarrolladas a lo largo de nuestra historia evolutiva (y cultural). Una extensión de este argumento señalaría que no es posible pasar de enunciados

descriptivos (sean sobre la evolución o sobre cualquier otra cosa) a enunciados normativos. Los argumentos que con más frecuencia se han esgrimido (sobre todo, desde la filosofía analítica) contra las éticas evolucionistas (y, en general, contra las éticas naturalizadas) son la denuncia de la falacia naturalista de Moore y la ley de Hume. Ello explica los esbozos de defensa frente a estos argumentos de autores como Casebeer, Hauser o Patricia Churchland. No puedo ocuparme en profundidad del análisis de la falacia naturalista y la ley de Hume, pero trataré de considerar y evaluar estos intentos de socavarlas o sortearlas.

A pesar de que ambos argumentos tengan relación y se hayan englobado frecuentemente bajo el mismo rótulo se trata de argumentos distintos, aunque ambos implicarían la autonomía de la ética frente a los hechos. Comencemos con Moore.

3. LA FALACIA NATURALISTA Y EL ARGUMENTO DE LA SINONIMIA

Moore llamó «falacia naturalista» a un tipo de error (en realidad, él admitió que no se trataba de una falacia lógica en sentido estricto) que cometen todas las éticas naturalizadas. Lo que Moore afirmó es, como hemos visto, que no es posible definir «bueno» en términos de alguna propiedad natural. Hay dos razones por las que esto es así: En primer lugar, las definiciones que le interesan a Moore no son las simples definiciones lexicográficas que indican el sentido usual de una palabra, sino las definiciones que aclaran la naturaleza real de algo. Pero este tipo de definiciones, afirma, «solo son posibles cuando el objeto o noción de que se trate sea algo complejo» (Moore, 1959: 6), porque consisten en indicar las partes que lo componen. «Bueno» no es una propiedad compleja, sino simple (porque, como amarillo, no puede descomponerse). Luego es indefinible. En segundo lugar, si alguien dijera: «Lo bueno es lo placentero» (o cualquier otra propiedad natural) estaría afirmando que «bueno» y «placentero» son sinónimos, de manera que una pregunta del tipo «¿Por qué lo placentero es bueno?» no tendría sentido. Sería como preguntar «¿Por qué lo bueno es bueno?», como sabría cualquier hablante competente. Sin embargo, la pregunta «¿Por qué lo pla-

centero (o cualquier otra propiedad) es bueno?» sí tiene sentido, es decir, siempre es una pregunta abierta (MOORE, 1959: 9 y ss.). En definitiva, MOORE trata de justificar una tesis ontológica (la existencia de dos tipos de propiedades: naturales y no naturales) a partir de un argumento semántico (el argumento de la pregunta abierta).

La argumentación de MOORE ha sido objeto de numerosas críticas. Una de las clásicas es la de FRANKENA. De acuerdo con este autor, hay tres posibles interpretaciones de la «falacia naturalista»: 1) Como el error consistente en definir una propiedad no natural («bueno») como si fuera natural, pero MOORE no demostró que «bueno» fuera una propiedad no natural, solo lo afirmó; 2) como el error consistente en definir una propiedad en términos de otra (obsérvese que si esta interpretación fuera correcta, entonces se estaría bloqueando todo análisis conceptual), pero para sostener esto MOORE debía haber mostrado en cada posible definición de «bueno» que era algo distinto de la propiedad con la que se emparejaba, cosa que no hizo; 3) como el error de tratar de definir lo simple e indefinible, pero tampoco MOORE demuestra que «bueno» sea una cualidad simple (solo lo afirma) (FRANKENA, 1974: 80; FOOT, 1974: 15). En definitiva, el ataque de MOORE parece basado en una petición de principio (o varias).

Como hemos visto, Patricia CHURCHLAND se defiende de la acusación de cometer la falacia naturalista de MOORE afirmando que, contra lo que MOORE parece pensar, cuando decimos «x es y» ambos términos no son necesariamente sinónimos. Si decimos que «el agua es H2O» no afirmamos que «agua» y «H2O» sean sinónimos. Si esto es así, entonces decir «El agua es H2O» no es equivalente a decir «El agua es agua» y preguntar por qué el agua es H2O no es una tautología, de manera que el argumento de la pregunta abierta no es correcto. Creo que en este punto CHURCHLAND tiene razón, como han reconocido también otros autores (PIGDEN, 2004: 574; MARTÍNEZ, 2003: 81). MOORE parece no estar distinguiendo entre sentido y referencia. Quien afirma que «lo bueno es lo placentero» afirma que ambos términos tienen la misma referencia, esto es, se refieren a la misma propiedad, aunque tengan sentidos o intensiones distintas. Parece haber ya cierto consenso en que MOORE no ha logrado justificar convenientemente la imposibilidad de identificar «bueno» con alguna propiedad natural

y, por tanto, su argumento no podría esgrimirse contra las teorías evolucionistas (en contra: RIVERA, 2002).

4. LA LEY DE HUME Y EL PRINCIPIO DE CONSERVACIÓN DE LA LÓGICA

En uno de los pasajes más citados de la historia de la ética (y que no me voy a resistir a citar otra vez) HUME afirmó, un poco de pasada, lo siguiente:

> En todo sistema moral del que haya tenido noticia, hasta ahora, he podido siempre observar que el autor sigue durante un cierto tiempo el modo de hablar ordinario, estableciendo la existencia de Dios o realizando observaciones sobre los quehaceres humanos, y, de pronto, me encuentro con la sorpresa de que, en vez de las cópulas habituales de las proposiciones: *es* y *no es*, no veo ninguna proposición que no esté conectada con un *debe* o *no debe*. Este cambio es imperceptible, pero resulta, sin embargo, de la mayor importancia. En efecto, en cuanto que este debe o no debe expresa alguna nueva relación o afirmación, es necesario que esta sea observada y explicada y que al mismo tiempo se dé razón de algo que parece absolutamente inconcebible, a saber: cómo es posible que esta nueva relación se deduzca de otras completamente diferentes. Pero como los autores no usan por lo común esta precaución, me atreveré a recomendarla a los lectores: estoy seguro de que una pequeña reflexión sobre esto subvertiría a todos los «sistemas corrientes de moralidad, haciéndonos ver que la distinción entre vicio y virtud ni está basada meramente en relaciones de objetos, ni es percibida por la razón» (HUME, 1981: 689 y 690).

La interpretación «estándar» de este pasaje (forjada en el marco de la filosofía analítica por autores como PRICHARD, AYER, HARE, NOWELL-SMITH o POPPER) entiende que HUME está señalando la imposibilidad de pasar de enunciados descriptivos a enunciados de deber ser, esto es, de derivar «lo que debe ser» a partir de afirmaciones sobre «lo que es», dado que se trata de tipos de proposiciones «enteramente diferentes». El efecto sería el mismo que el producido por el argumento de la falacia naturalista: la ética (la normatividad, en general) quedaría como un ámbito autónomo, cuya fundamentación

—si la tiene— no puede encontrarse en los hechos. Tan solo esta razón ya haría imposible una ética naturalista. Pero esta interpretación no está exenta de problemas y puede resultar sospechosa de no tener en cuenta el contexto en el que HUME escribió. Para empezar, porque el propio HUME parece haber construido una ética naturalista, basada en las emociones (y además con varios puntos de coincidencia con las éticas evolucionistas). Es decir, el propio HUME estaría incumpliendo la ley de Hume si esta lectura fuera correcta (MACINTYRE, 1959: 451-468; CHURCHLAND, 2012: cap. 1). Para preservar su coherencia debería buscarse otra interpretación. Por ejemplo, el final del texto transcrito (que, de hecho, muchas veces se omite en las reproducciones del pasaje) podría servir para sugerir que lo que HUME pretende no es desconectar la moral (el vicio y la virtud, diría él) de los hechos, sino señalar que esa conexión no puede ser establecida por la razón, es decir, no puede fundarse racionalmente. La conexión vendría establecida por nuestras reacciones de agrado o desagrado (guiadas por la «simpatía») frente a ciertas acciones, por nuestras pasiones, sentimientos, deseos o intereses.

Otro problema que plantea la interpretación del texto de HUME es la afirmación de que el paso del ser al deber ser «parece absolutamente inconcebible»: ¿Quiere decir, efectivamente, que es imposible? (interpretación estándar) ¿Que no es fácil concebir cómo hacerlo (esto es, que *parece* imposible, pero sin llegar a afirmar que lo sea)? (MACINTYRE, en la obra citada). ¿O lo que es inconcebible es que los autores anteriores nunca hayan explicitado cómo realizan el paso de proposiciones de un tipo al otro? (DUSELL, 2001: 67). Por otra parte, también la expresión «deducir» es problemática: ¿se refiere HUME a una deducción en sentido estricto o usa la expresión en un sentido amplio (como equivalente a inferencia)? ¿Y aceptaba HUME la validez de otros tipos de inferencias o las consideraba «defectuosas»? La interpretación estándar señala que HUME usa la expresión «deducir» en sentido estricto y, además, que consideraba «defectuosas» otras formas de inferencia (por lo que no hay otra vía racional para pasar de hechos a normas), pero algunos autores han señalado que en el contexto de su época «deducir» se entiende mejor como «inferir» (MACINTYRE).

Pero podemos dejar de lado estas dudas porque, en realidad, la cuestión de qué quiso decir de hecho HUME es secundaria para la

validez de lo que conocemos como ley de Hume, esto es, para la cuestión de si se puede pasar o no del «ser» al «deber ser». Lo relevante es si hay alguna interpretación correcta de la tesis, no si HUME la sostuvo o no. Una manera muy simple de defender la imposibilidad de pasar de enunciados descriptivos a enunciados normativos (o valorativos) consiste en relacionar esta cuestión con el principio de conservación de la lógica: en una deducción no es posible que se concluya algo que no estuviera incluido ya en las premisas. Las deducciones pueden hacer que seamos *conscientes* de un dato nuevo, pero este se encontraba ya en las premisas. Por ello, solo de proposiciones descriptivas no podemos deducir enunciados de deber ser. Esto ocurre con cualquier cosa. Como observa PIGDEN «existe un salto similar entre las conclusiones sobre los erizos y las premisas que no hacen mención de ellos. No se pueden obtener conclusiones sobre "erizos· a partir de premisas carentes de erizos (al menos no solo por la lógica)»[1] (PIGDEN, 2004: 570). Si aceptamos esto, entonces es obvio que muchos de los argumentos en los que aparentemente pasamos de enunciados descriptivos a enunciados prescriptivos son en realidad argumentos deductivos entimemáticos, en el sentido de que tienen una premisa implícita prescriptiva o valorativa. Esta es una manera de entender el argumento con el que HAUSER cree haber mostrado la falsedad de la ley de Hume:

> Hecho: la única diferencia entre un médico que aplica anestesia a una criatura y otro que no se lo aplica es que, sin ella, el niño sufrirá enormemente en el curso de una operación. La anestesia no tendrá ningún efecto pernicioso para el niño, sino que le hará perder

[1] PIGDEN usa esta idea para convertir la tesis de HUME en algo trivial y sin consecuencias para las éticas naturalistas: opina que del hecho de que las normas morales no puedan deducirse de afirmaciones se sigue la autonomía lógica de la moral, pero no su autonomía o independencia ontológica, que es lo que neceistan las éticas naturalistas. De la misma manera, del hecho de que no podamos inferir nada sobre los erizos de premisas que no mencionan erizos se sigue su autonomía lógica, pero no su autonomía ontológica (su pertenencia a un «reino de los erizos» separado y sin relaciones con cualquier otro). «Nadie —escribe— propone una distinción hecho/erizo, o alega que las proposiciones sobre los erizos no son en realidad proposiciones sino una singular subclase de madnatos» (PIGDEN, 2004: 570). Ahora bien, no estoy seguro de que PIGDEN esté en lo cierto cuando piensa que la autonomía lógica de las normas no afecta a las éticas naturalistas, porque lo que estas pretenden ens que pueden justificar qué debemos hacer a partir de afirmaciones sobre el mundo, y la relación de justificación sí requiere inferencias.

> temporalmente la conciencia y la sensación de dolor. Luego despertará, una vez acabada la operación, sin ninguna secuela negativa y con mejor salud, gracias al trabajo del médico.
>
> Juicio valorativo: por consiguiente, el médico debe administrar anestesia al niño (HAUSER, 2008: 28).

Si el deber de la conclusión es un deber deóntico (aunque HAUSER lo llame «juicio valorativo»), entonces puede reconstruirse este argumento como una inferencia deductiva, cuya validez presupone una «premisa mayor» según la cual: «Debe evitarse el sufrimiento innecesario del niño» (y, quizá, alguna premisa adicional más). Luego este argumento (y todas las derivaciones aparentes que en realidad se basan en premisas implícitas) no cuenta como objeción a la ley de Hume.

Otras veces, lo que ocurre no es tanto que el argumento que concluye un deber sea entimemático como que incluya una premisa ambigua, que aparentemente sea descriptiva, pero en realidad no lo sea. Eso es probablemente lo que ocurre en uno de los intentos más conocidos de derivación de «debe» a partir de «es»: el de John SEARLE (aunque no es una vía que haya seguido la neuroética, seguramente porque depende de contextos institucionales, y no «naturales»). Lo que SEARLE trata de demostrar puede resumirse de la siguiente manera:

a) La representación clásica del mundo no logra dar cuenta de los hechos institucionales.

b) Los hechos institucionales existen dentro de sistemas de reglas constitutivas.

c) Algunos sistemas de reglas constitutivas incluyen obligaciones, compromisos y responsabilidades.

d) Dentro de algunos de estos sistemas podemos derivar «debe» a partir de «es».

El argumento de SEARLE para demostrar *d)* consiste en tratar de concluir un juicio *normativo* moral acerca de un sujeto partiendo de un juicio que *describe* una cierta forma de conducta de ese sujeto, a través de premisas intermedias de índole igualmente descriptiva. Para ello utiliza la institución de la promesa. El razonamiento sería el siguiente (SEARLE, 1980: 178-201):

1) *A* ha pronunciado las palabras siguientes: «Prometo pagar a *B* mil euros».
2) *A* ha prometido pagar a *B* mil euros.
3) *A* se ha obligado a pagar a *B* mil euros.
4) *A* tiene la obligación de pagar a *B* mil euros.
5) *A* debe pagar a *B* mil euros.

La relación entre cualquiera de estos enunciados y el siguiente, «aunque no es en todo caso —escribe SEARLE— una relación de entrañamiento, no es sin embargo una relación puramente accidental o completamente contingente» (SEARLE, 1980: 180). El argumento, sin embargo, ha sido criticado en numerosas ocasiones y no ha generado aceptación. Una crítica común sugiere que debe tenerse en cuenta la distinción entre asumir un punto de vista interno o un punto de vista externo frente a una institución. Si nos colocamos en un punto de vista externo a una institución (por ejemplo, una promesa), solo estamos informando de la existencia de la misma, por lo que nos situamos en un nivel descriptivo; pero si nos situamos en un punto de vista interno, estamos asumiendo una posición *comprometida* y normativa, con lo que pasamos a un nivel prescriptivo o valorativo. Lo que puede criticarse en el argumento de SEARLE es el paso inadvertido de un punto de vista externo a un punto de vista interno. Solo si el sujeto acepta la institución de la promesa (y entonces introduce un componente normativo en las premisas) se llega a una conclusión normativa [puede verse una síntesis de las críticas a SEARLE en HUDSON, 1987: 269 y ss.; HIERRO, 1970: 143-144; GONZÁLEZ LAGIER, 1995: 273 y ss.; GUASTINI, 1990: 272 y ss.][2].

¿Hay alguna manera de escapar de esta forma de entender la ley de Hume? Examinaré dos posibles estrategias: 1) sostener que no siempre es posible reconstruir los argumentos en los que parece derivarse «debe» de «es» como deducciones entimemáticas, o

[2] Por su parte, VON WRIGHT sugiere tres interpretaciones (que evitarían el «salto» del «ser» al «deber» en el argumento de SEARLE): o bien se trata de un argumento entimemático cuyas premisas deben completarse con una norma implícita según la cual «No debe ser el caso que alguien haga una promesa sin cumplirla»; o bien la conclusión no es una norma sino una proposición normativa (descriptiva, por tanto); o bien la conclusión es un deber técnico acerca de qué tenemos que hacer si queremos cumplir con nuestra promesa, con lo que tampoco obtenemos un deber genuino [VON WRIGHT, 1985: 273-275; GONZÁLEZ LAGIER, 1995: 273 y ss].

2) sostener que es posible derivar «debe» de «es» por medio de argumentos a los que no se aplica el principio de conservación, como la inducción o la abducción.

5. LA LEY DE HUME Y LA DERROTABILIDAD DE LAS NORMAS

La primera estrategia ha sido sugerida por Ulises MOULINES. En un artículo sumamente recomendable (sobre todo por su llamada de atención sobre la errada interpretación de muchas dicotomías conceptuales como excluyentes), dedicado a la dicotomía hecho/valor, MOULINES pone el siguiente diálogo como ejemplo de derivación de «debe» (más exactamente, de una valoración moral) a partir de «es» (MOULINES, 1991: 39):

A.—Luis estranguló a su mujer.

B.—¡Cómo! ¿Se volvió loco?

A.—No; aunque eso es lo que le está tratando de hacer creer a la policía.

B.—Pero, entonces, ¿fue un acto irreflexivo, debido a un momento de ofuscamiento?

A.—No; todo indica que lo planeó desde hace tiempo; estaba dando a entender a la gente que su mujer se había vuelto loca y que se iba a suicidar, precisamente estrangulándose a sí misma. Además, pensó que, aunque lo descubrieran, no le iba a pasar nada; es un hombre con muchas influencias, amigo del ministro X.

B.—Pero, entonces, ¿qué motivo tuvo?

A.—Se quería ir a vivir con su amante.

B.—Entonces, este Luis es un canalla, un malvado.

Cualquier persona, dice MOULINES, reconocería este argumento como adecuado. Podría pensarse que su validez depende de una premisa implícita, que vendría a decir: «Todo hombre que se hace pasar por loco a sí mismo y a su mujer, que tiene influencias, que es amigo de ministros, tiene una amante, y estrangula a su mujer, es un malvado». Esta premisa convertiría el argumento en un simple *modus ponens* del que se deduciría que Luis es un malvado. MOULINES señala que sería extraño pensar que esta premisa forma parte

de algún sistema moral, pero cabría decir que se deriva o se construye a partir de otras normas del sistema. El problema principal, sin embargo, sería que es fácil añadir información que haría que la conclusión no se siguiera:

> No vayamos tan deprisa en juzgar a Luis. Es cierto que estranguló a su mujer y que está haciéndose el loco sin estarlo realmente ahora; pero sabemos que con frecuencia antes tuvo brotes de paranoia, y ciertos detalles indican que se hallaba en tal estado cuando estranguló a su mujer; además es verdad que quería irse a vivir con su amante, pero un día antes de que estrangulara a su mujer, rompió con ella. Y finalmente no hay que olvidar que su mujer constantemente lo amenazaba, incluso físicamente, y es posible que esto ocurriera esa noche». ¿Qué concluirá ahora B? Quizá: «Este Luis es un desdichado». En cualquier caso no concluirá «Luis es un malvado» (MOULINES, 1990: 40).

Ahora, la premisa normativa que puede dar «cobertura» al razonamiento que hay detrás del anterior razonamiento es más compleja, porque debe dar cuenta de las circunstancias anteriores y de las nuevas, esto es, de los hechos «constitutivos» de la calificación moral «malvado» y de los hechos que la cancelan para el caso concreto. Pero podrían aportarse nuevos datos que inclinaran la balanza de nuevo hacia la calificación de malvado, y habría de nuevo que revisar y modificar la norma de cobertura, y así hasta el infinito. La conclusión de MOULINES es que: «Ninguna situación humana (o empírica en general) puede describirse de manera tan completa que sea base suficiente para una premisa moral universal del tipo de las que necesita el dicotomista para fundamentar una conclusión moral particular sobre esa situación. Ese "programa", si así puede llamarse, es simplemente irrealizable. Pero eso significaría, para el dicotomista, que en la inmensa mayoría de casos en los que hacemos juicios de valor, él debería abstenerse de hacerlos» (MOULINES, 1990: 41). Es decir, no sería aceptable considerar estos argumentos como entimemáticos porque la premisa normativa sería imposible de formular.

El problema al que MOULINES está apuntando es, obviamente, el de la derrotabilidad de las normas, más exactamente, del razonamiento normativo. Los teóricos del derecho suelen aceptar esta

característica del razonamiento, al menos desde que HART señalara que es imposible dar una lista de todas las condiciones necesarias y suficientes para la aplicación de los conceptos jurídicos (y morales, cabría añadir: la noción de David ROSS de *deber prima facie* apunta en la misma dirección). Y la actual discusión acerca de reglas y principios ha generado una mayor profundización en la discusión sobre la derrotabilidad. El problema, además, no es exclusivo de los razonamientos normativos. Como ha señalado Juan Carlos BAYÓN: «en la vida cotidiana extraemos intuitivamente conclusiones a partir de generalizaciones que tienen excepciones —no enumerables exhaustivamente de antemano—, porque se refieren a propiedades *típicas* o *normales* —en relación a cierto contexto— de individuos en un dominio» (BAYÓN, 2000: 92), es decir, obtenemos conclusiones a partir de generalizaciones que admiten excepciones no enumerables. Todo ello plantea un problema lógico, ya que la lógica clásica es monotónica, esto es, si una conclusión se sigue de un conjunto de enunciados *E*, entonces se sigue de cualquier otro conjunto de enunciados que incluya todos los elementos de *E*. Algunos autores consideran que para dar cuenta de razonamientos derrotables (no monotónicos) habría que construir una lógica no monotónica e introducir la idea de condicionales derrotables (tanto en la lógica deóntica, si se quiere dar cuenta del razonamiento normativo, como en la lógica proposicional, si se quiere dar cuenta del razonamiento con generalizaciones). Pero, entonces, no queda claro cómo hacer esto sin poner en juego la regla del *modus ponens* y la idea misma de deducción. Otros autores no creen que sea necesario. Por ejemplo, ALCHOURRÓN sostiene que: «No es necesario una lógica para normas derrotables (normas de obligaciones y permisiones *prima facie*) porque detrás de la demanda por esta clase de lógicas, así como detrás de la demanda por lógicas no monotónicas subyace una combinación de una noción ordinaria de consecuencia (o condicional) y el cambio de nuestras premisas (normativas) en una perspectiva dinámica» (ALCHOURRÓN, 2010: 78; MORESO; RODRÍGUEZ, 2010: 15 y ss].

Todo esto es conocido, y además se trata de un problema abierto en el que no es posible profundizar aquí. Pero, ¿deberíamos extraer de ahí que la lógica deductiva no puede decir nada sobre cuestiones normativas y empíricas y queda recluida exclusivamente al ámbi-

to formal? Creo que no. MOULINES afirma que: «Ninguna situación humana (o empírica en general) puede describirse de manera tan completa que sea base suficiente para una premisa moral universal del tipo de las que necesita el dicotomista para fundamentar una conclusión moral particular sobre esa situación», pero para justificar la conclusión (una justificación revisable) no necesitamos una premisa completa. Basta con aceptar que esta pueda ser revisada. Y, como ha escrito BAYÓN: «no hay que entender que esta inferencia sea algo distinto de la deducción clásica solo porque su conclusión decaiga cuando una de sus premisas haya de ser reemplazada porque se ha revelado errónea. Que la premisa sea "derrotable" (en el sentido de falible o revisable) no implica que la inferencia misma sea derrotable (en el sentido de no deductiva)» (BAYÓN, 2001: 36). El argumento de MOULINES parece incurrir en una falacia que podríamos llamar «falacia absolutista»: se comienza por exigir que la norma que pretende justificar una decisión sea completa y definitiva, se constata que es imposible y se concluye que la justificación no depende de la norma. Por la misma razón, como no podemos tener nunca una prueba definitiva de que nuestras creencias estén justificadas, habría que concluir que no debemos tratar de fundamentar nuestras creencias en pruebas (por supuesto, con esto no quiero decir que la lógica deductiva sea suficiente para dar cuenta de toda la complejidad de los razonamientos normativos o empíricos reales).

Los argumentos normativos entimemáticos sí pueden, por tanto, ser completados (con premisas revisables). Por tanto, por esta vía no eludimos la ley de Hume.

6. LA LEY DE HUME Y LOS ARGUMENTOS NO DEDUCTIVOS

La segunda estrategia es la seguida por CASEBEER y por P. CHURCHLAND. Para CASEBEER la ley de Hume veta solo inferencias deductivas, pero no inferencias a la mejor explicación (CASEBEER, 2003: 482). Y CHURCHLAND afirma: «Desde la perspectiva de la neurociencia y la evolución cerebral, el rechazo rutinario de los enfoques científicos de la conducta moral basada en la advertencia de Hume en contra de la derivación de normas a partir de hechos pa-

rece desafortunado, especialmente porque la advertencia se limita a las inferencias deductivas» (CHURCHLAND, 2012: cap. 1). Para esta autora, el pasaje de HUME debe ser interpretado como una advertencia contra las inferencias simplistas, no como un veto a cualquier tipo de paso del ser al deber ser.

A diferencia de las deducciones, las inducciones y las abducciones sí amplían nuestro conocimiento, por lo que no rige para ellas el principio de conservación. Por tanto, si la ley de Hume es solo una manifestación del principio de conservación de los argumentos deductivos, entonces tienen razón los que sostienen que esta no es aplicable a las inferencias no deductivas. Pero, ¿es solo eso la ley de Hume? Probablemente, no.

Puede sostenerse que entre enunciados descriptivos y enunciados normativos existen importantes diferencias —a veces se habla de un «abismo lógico» entre ellas, o entre hechos, por un lado, y normas y valores, por otro: así, los enunciados descriptivos tienen una dirección de ajuste descendente (esto es, palabras a mundo: se pretende que las palabras se ajusten al mundo) mientras que los enunciados normativos tienen una dirección de ajuste ascendente (esto es, mundo a palabras: se pretende que el mundo se ajuste a las palabras)—. Los enunciados descriptivos son verdaderos o falsos, mientras que las normas o valores no lo son. Los enunciados que expresan deberes presuponen un punto de vista interno, en el que los términos deónticos (obligatorio, prohibido) son *usados*, mientras que, si una descripción se refiere a una norma o a un deber, lo hace desde un punto de vista externo en el que los términos deónticos son solo mencionados. Todas estas diferencias entre descripciones y normas hacen que *las primeras no puedan servir de razones, ni expresar razones, para justificar las segundas*. No es solo que una justificación deductiva requiera que entre las premisas esté aquello que se quiere deducir, es que —aun cuando se admita que la inducción o la abducción puedan tener alcance justificatorio— ningún enunciado descriptivo puede, por sí solo, ser una razón que justifique un enunciado prescriptivo. Puede aportar una razón explicativa de por qué aceptamos ciertas normas o valores; puede ser también una razón explicativa o, incluso, justificatoria, de otros enunciados descriptivos. Pero nada más. Todo esto, por cierto, requiere ciertas presuposiciones (NAVARRO, 2007: 602 y ss). Por ejemplo, si se con-

sidera que las normas son verdaderas o falsas, este «abismo lógico» desaparece o se estrecha, pero entonces el defensor de esta concepción de las normas debe afrontar importantes dificultades.

Pensemos en la forma de las inducciones y las abducciones. Las inducciones consisten en pasar a partir del examen de varios ejemplos a formular un enunciado general. En este caso, ese enunciado general sería una norma. Pero entonces lo que tenemos es el tipo de inferencias simplistas que descarta CHURCHLAND. Compárense los siguientes argumentos:

1) *X* aspira a la felicidad.
Y aspira a la felicidad.
Z aspira aspira a la felicidad.

Luego no debemos obstruir la búsqueda de la felicidad de los demás.

2) En los animales de la especie *x* los fuertes eliminan a los débiles.
En los animales de la especie *y* los fuertes eliminan a los débiles.
En los animales de la especie *z* los fuertes eliminan a los débiles.
Luego debemos permitir que los fuertes eliminen a los débiles.

Supongamos que estamos de acuerdo con la conclusión del primer argumento y no del segundo. ¿cómo explicaríamos esto? Podremos decir que es porque las premisas del primer argumento son adecuadas para inferir una conclusión de naturaleza moral, pero las segundas no. ¿Pero por qué pensamos así? Lo pensamos porque evaluamos estos argumentos con premisas normativas o valorativas. Luego estos argumentos que parten de premisas descriptivas no justifican por sí solos la conclusión normativa a la que pretenden llegar. Necesitamos criterios normativos o evaluativos para seleccionar qué hechos tomamos como (parte de las) premisas de nuestros razonamientos normativos.

Pensemos ahora en la abducción. Podemos usar la abducción para establecer que es razonable aceptar una hipótesis o teoría, cuando la verdad de esa hipótesis explicaría un hecho que nos parece sorprendente. Este tipo de abducción tendría la siguiente estructura:

Se observa el hecho sorprendente, *F*.
Pero si *H* fuera verdadero, *F* sería cosa corriente.
Por lo tanto, hay razón para sospechar que *H* es verdadero.

¿Cómo podría usarse este argumento para pasar de enunciados descriptivos a normas o valores? ¿Habría que asumir que lo que mejor explica ciertos comportamientos es la existencia de una regla o un deber moral? Esto lo único que haría —a lo sumo— sería explicar una serie de comportamientos en la creencia de que existe un deber (creencia que es otro hecho), pero no justificaría para nada ese deber. Hagamos otro intento: FLANAGAN dice que el buen naturalista ético debe mostrar que ciertas normas o valores son razonables por medio de la abducción (FLANAGAN, 1996: 22). Supongo que el razonamiento sería el siguiente:

Una norma moral razonable tendría las consecuencias *x*, *y* y *z*.
La norma N tiene las consecuencias *x*, *y* y *z*.
Luego la norma N es razonable.

Pero parece evidente que la primera premisa presupone un juicio de valor (las consecuencias *x*, *y* y *z* se consideran valiosas). El mundo es axiológicamente neutro y de él, sin proyectar nuestros valores, no se deriva nada[3].

Me parece, además, que en buena medida el rechazo del papel de la deducción en el razonamiento moral se debe a una falta de distinción entre explicación del modo como resolvemos problemas

---

[3] ¿Pero es realmente el mundo axiológicamente neutro? Uno de los argumentos que podría debilitar la ley de Hume es la idea de que los hechos (incluso los descritos por la ciencia) «están cargados de valores» o «cargados de normas». Este sería el argumento de PUTNAM. En el caso del derecho podríamos pensar de la siguiente manera: 1) La omisión es un tipo de hecho; 2) «*X* omitió ayudar a *y*» es una descripción; 3) de esa descripción podemos derivar «*X* tenía el deber de ayudar a *y*». Esto es así porque, por definición, existe una omisión solo si se ha incumplido un deber. De esta manera, derivamos normas o deberes de enunciados descriptivos. Creo que debemos aceptar la existencia de hechos normativos en este sentido (hechos cuya definición remite a una norma). Ahora bien, esa misma remisión hace que en el hecho (y en su descripción) ya haya un componente normativo. Lo que permite derivar el deber es ese componente normativo. Creo que puede aceptarse que muchos hechos (entendidos no como datos brutos de la realidad, sino como hechos interpretados a la luz de reglas y convenciones sociales) son amalgamas de eventos brutos y convenciones (GONZÁLEZ LAGIER, 2007), sin que ello ponga en peligro la ley de Hume. Creo que la cuestión es semejante al ejemplo de la promesa de HUME.

morales y justificación de los mismos (o, si se quiere, entre el contexto de descubrimiento y el de justificación). Por ejemplo, Patricia CHURCHLAND sostiene que «el razonamiento práctico consiste en encontrar una buena solución a un problema de satisfacción de restricciones. La deducción —el favorito sentimental de los lógicos— juega como mucho un papel menor o *post hoc*» (CHURCHLAND, 2008)[4]. Lo que sugiere es que los problemas sociales y morales implican «diversos factores de distinta importancia y probabilidades que interaccionan hasta derivar en una solución óptima a una cuestión» (CHURCHLAND, 2012: cap. 1). Ahora bien, esto puede ser adecuado para explicar el proceso de toma de decisiones (y de ahí el interés de este tipo de problemas para la inteligencia artificial), pero no es suficiente para establecer si estas están justificadas.

7. LA LEY DE HUME Y LOS DEBERES TÉCNICOS O PRUDENCIALES

No siempre que un enunciado incluye el término «deber» es un genuino enunciado normativo. A veces, «debe» expresa una conjetura («debe ser así» puede significar «probablemente es así»), otras veces, se puede sustituir por «tiene que» y expresa una necesidad práctica. Es importante distinguir entre deberes deónticos o genuinos y deberes técnicos, prudenciales o necesidades prácticas. Muchos de los ejemplos que se ofrecen como derivación de «debe» a partir de «es» no concluyen genuinos deberes deónticos, sino necesidades prácticas. Como señala VON WRIGHT, podemos «encontrar dos respuestas principales a la cuestión de por qué una cierta cosa debe o puede o no tiene que ser hecha. Una es que existe una norma ordenando o permitiendo o prohibiendo la realización de esa cosa. La otra es decir que los fines y las conexiones necesarias hacen (o no) la realización u omisión de esa cosa una necesidad práctica»

---

[4] Un «problema de satisfacción de restricciones» (*Constraint Satisfaction Problem*) es un tipo de problema donde la solución (o soluciones) debe satisfacer un conjunto de requisitos o restricciones. Por ejemplo colorear los países de un mapa de manera que no se coloreen dos países adjuntos con el mismo color, o un sudoku, o colocar *n* reinas en un tablero de ajedrez de *nxn* casillas de manera que no se ataquen (MANYA y GOMES, 2003).

(VON WRIGHT, 1983: 74). En *The Varieties of Goodness* VON WRIGHT define las necesidades prácticas como aquellas que surgen de la combinación del deseo de un determinado fin con el conocimiento de los medios necesarios para alcanzarlo. Deseo y creencia se combinan en un silogismo práctico con la siguiente forma:

a) Tu quieres q.
b) A menos que hagas p, no conseguirás q.
c) Por tanto, tienes que hacer p.

La primera premisa expresa un deseo; la segunda premisa expresa una necesidad (causal, si depende de una ley de la naturaleza, o convencional, si depende de una institución o una norma prescriptiva); por último, la conclusión expresa lo que VON WRIGHT propone llamar una necesidad práctica (VON WRIGHT, 1963: 161). Los deberes deónticos o genuinamente normativos, a diferencia de las necesidades prácticas, son —pretenden ser— por sí mismos razones operativas para la acción, esto es, razones que compelen a actuar. Las necesidades prácticas solo mueven a la acción en la medida en que se desee el fin. La razón operativa es, entonces, el deseo del fin. Por otra parte, los deberes deónticos no son verdaderos o falsos, mientras que las necesidades prácticas sí lo son (en relación con el fin deseado) (VON WRIGHT, 1976: 85).

Uno de los intentos de derivación del deber ser a partir del ser se debe a Max BLACK, quien toma, precisamente, a los juicios normativos técnicos y a la inferencia práctica como eje de su argumentación. El ejemplo de Max BLACK es el siguiente:

a) *A* quiere dar jaque mate a *B*.
b) A menos que *A* mueva la reina no podrá dar jaque mate a *B*.
c) *A* debe mover la reina.

Tanto *a)* como *b)* son premisas descriptivas. La conclusión *c)* es, sin embargo, según este autor, una premisa prescriptiva. Que se use como ejemplo el juego del ajedrez puede provocar el error de suponer que se trata de una argumentación semejante a la de John SEARLE, pero lo relevante aquí no es que se trate de un contexto institucional (el juego, en este caso; la promesa, en el intento de SEARLE): si *A* debe mover la reina no es porque ése sea un movimiento necesario para jugar al ajedrez —seguiría jugando aun cuando no

lo hiciera—, sino porque es la única manera de dar jaque mate en esa situación. En contra de BLACK, VON WRIGHT niega que nos encontremos ante un caso en el que se haya producido un paso del ser al deber ser. La conclusión, para VON WRIGHT, es un deber técnico o necesidad práctica, que pertenece al mundo del ser (VON WRIGHT, 1985: 265 y ss.).

Es interesante advertir que estos deberes técnicos pueden aparecer enlazados a deberes genuinamente normativos, lo que puede hacer difícil distinguirlos. Dado un deber deóntico, puede ser que consideremos que, a menos que hagamos *x* o *y*, no podremos cumplir con ese deber; entonces surge un deber técnico, una necesidad práctica, de la existencia de un deber deóntico [VON WRIGHT, 1985: 276 y 277]. Si consideramos la posibilidad de que los deberes deónticos den lugar a deberes técnicos, entonces es fácil confundir unos y otros.

Pues bien, los ejemplos de la vida cotidiana que Patricia CHURCHLAND aporta como derivaciones de deberes a partir de enunciados fácticos son, en realidad, argumentos que se mueven exclusivamente en el plano descriptivo o que concluyen en necesidades prácticas, no en deberes genuinos: «Tengo un dolor de muelas espantoso? Entonces debería ir al dentista. ¿El horno está ardiendo? Debería apagarlo con bicarbonato de soda. ¿Hay un oso en mi camino? Debería caminar despacio, tarareando para mis adentros, siguiendo una trayectoria octogonal» (CHURCHLAND, 2012: cap. 1) (también el ejemplo de HAUSER que hemos visto más arriba puede interpretarse como concluyendo una necesidad práctica).

Por supuesto, razonamientos de este tipo pueden darse también sobre cuestiones de relevancia social y moral. Por ejemplo, «si queremos evitar la degradación de nuestro planeta debemos reducir la emisión de gases con efecto invernadero». Esta es una regla prudencial o técnica que se deriva del deber de evitar la degradación de nuestro planeta. Este último deber puede verse a su vez como un deber técnico derivado del deber de evitar la extinción de la humanidad. Pero en algún momento esta cadena de deberes técnicos debe cerrarse con un deseo último o con un deber deóntico. Solo si se cierra con un deber deóntico podemos hablar genuinamente de normatividad. Es sugerente pensar que los sistemas morales son un

conjunto de reglas técnicas en relación con una serie de fines que se toman como axiomas, pero lo que les confiere a esos sistemas su carácter moral es estar supeditados a fines que se consideran moralmente valiosos (y no solo deseados), esto es, deberes morales últimos, que ya no pueden ser fundamentados en un hecho.

Podría pensarse, en definitiva, que la ley de Hume no impide derivar enunciados de deber (en sentido amplio) a partir de enunciados sobre hechos, sino solo genuinos deberes a partir de enunciados fácticos. Es decir, el mundo del «deber ser» abarca entidades solo unidas por cierto parecido de familia y a la que por costumbre llamamos «normas» a veces de manera imprecisa (von Wright, 1970: 23 y ss.). La llamada «gran división» entre normas y descripciones no se da entre todas aquellas cosas que realmente llamamos normas, sino entre enunciados descriptivos, por un lado, y genuinos enunciados prescriptivos, por otro.

EPÍLOGO

La neuroética (en el sentido que se le ha dado aquí) es el intento de iluminar el razonamiento (aunque acaba concluyendo que se trata más bien de un no-razonamiento) y el comportamiento moral a la luz de los conocimientos neurocientíficos. En un primer nivel, un nivel descriptivo, la neuroética sostiene que nuestras opiniones morales dependen de una facultad, la intuición (frecuentemente conectada o identificada con las emociones) que nos da la respuesta al dilema moral que hemos de resolver. Sostiene, además, que parte de esas intuiciones son innatas y se han desarrollado de acuerdo con las leyes de la evolución, porque valores como la reciprocidad, la cooperación o el altruismo favorecen la supervivencia de la especie. Ya hemos visto que estas afirmaciones deben verse como hipótesis especulativas. Se basan, ciertamente, en pruebas, pero —en mi opinión— no se han descartado hipótesis alternativas (que las intuiciones deriven de la razón, que los supuestos conocimientos innatos sean en realidad culturales, que la evolución haya favorecido la inteligencia, pero no directamente nuestras capacidades morales y, mucho menos, principios o valores éticos...) que impiden aceptar definitivamente las primeras. Se trata de una tarea realmente interesante, pero en la que es mejor avanzar con prudencia, como la que exhibe en España Camilo J. CELA CONDE ante el abanico de interrogantes abiertos:

> ¿Estamos en el caso del juicio moral o de otros fenómenos perceptivos similares ante procesos cognitivos más bien unitarios y discretos, o se trata solo de fenómenos que emergen de muchos mecanismos psíquicos articulados en el tiempo y el espacio? ¿Tienen

> esos presuntos procesos o series de procesos algún aspecto de carácter universal, en el sentido de que cuenten con alguna componente clave común capaz de determinar en cada individuo su particular valoración de lo que es o deja de ser justo? ¿Será posible algún día describir ese proceso o procesos (o las componentes clave) en términos más objetivos? ¿Cabe buscar su origen en algún patrón idiosincrásico de actividad neural que contenga al menos alguna secuencia espacio-temporalmente identificable compartida por todos los sujetos? [...] ¿En qué medida contribuyen la herencia y la historia de aprendizaje de cada individuo en la puesta en marcha de ese supuesto patrón funcional? (CELA *et al.*, 2005: 132).

En un segundo nivel, un nivel normativo, algunos autores conceden relevancia normativa a las intuiciones morales que se postulan en el primer nivel. Las intuiciones nos dicen qué debemos hacer, y dado que esas intuiciones vienen con el aval de nuestro éxito evolutivo, lo correcto es hacer lo que ellas nos dicen (o bien estamos determinados a hacerlo así, en cuyo caso la normatividad, sencillamente, desaparece: queda desplazada por la causalidad). Una propuesta normativa de este tipo no puede hacerse sin presuponer una metaética (la metaética tiene, ciertamente, una dimensión conceptual importante, pero puede ir más allá y tener una pretensión justificatoria). No es fácil decir qué metaética está detrás de la neuroética prescriptiva, porque los autores que la han propuesto no han tratado de articularla, porque todavía es una teoría joven (aunque puede verse como una rama del evolucionismo moral) y porque integra elementos de varios tipos de teorías. Así que clasificarla es en parte un ejercicio imaginativo (en el que hay que suponer algunas cosas aún no claras o consensuadas entre los autores). Pero creo que podría calificarse la metaética detrás de la neuroética con pretensiones normativas (del «realismo moral neurocientífico») de ser —desde un punto de vista filosófico— una extraña combinación: un «naturalismo ético intuicionista y emotivista». Trataré de justificar esta calificación.

De acuerdo con RIVERA LÓPEZ, el naturalismo ético tiene las siguientes características (RIVERA, 2002: 52):

a) Es descriptivista: entiende que los juicios morales del tipo «esta es la decisión correcta» o «esto es bueno» son verdaderos o falsos.

b) Es realista: los juicios morales verdaderos lo son por su referencia a ciertos hechos.

c) Es reduccionista: los hechos que hacen verdaderos o falsos a los juicios morales son hechos no-morales.

d) Puede ser «empirista» (si consideran que los hechos que hacen verdaderos o falsos a los juicios morales no son de carácter metafísico) o no empirista (si conciben tales hechos como metafísicos: por ejemplo, la existencia de Dios, un Derecho natural, etc.).

Pues bien, la neuroética normativa es descriptivista y realista: considera que hay ciertos hechos de los que depende que lo que hagamos o decidamos sea correcto o no (vendría a ser algo así como que lo correcto es lo que nos dicen nuestras intuiciones avaladas por la evolución). Es reduccionista, puesto que tratan de reducir las propiedades morales a propiedades no morales (creencias, intuiciones, emociones... determinadas por el proceso de evolución). Y es empirista (tales propiedades no morales pueden ser analizadas por medio de la ciencia). Estas características permiten afirmar que es un tipo de naturalismo ético. Sin embargo, la neuroética normativa tendría dos ingredientes que toma de teorías rivales del naturalismo: por un lado, comparte con el intuicionismo ético de los filósofos la idea de que los juicios morales no se derivan de la razón, sino de la intuición. La intuición es la que nos dice si una determinada acción es correcta o no, aunque de aquí no seguiría para estos autores que las propiedades morales no sean naturales. Por otro lado, comparte con el emotivismo la idea de que los juicios morales son expresiones de nuestras emociones (en la medida en que identifica intuiciones y emociones), aunque se distingue del emotivismo clásico en que ello no le lleva necesariamente a sostener posturas escépticas acerca de la objetividad de la moral.

No obstante, creo que el proyecto de una neuroética normativa, en la medida en que pretenda derivar normas morales sin fundamentar esta pretensión en algún principio normativo último, es inviable. No es posible justificar normas, valores o principios exclusivamente en hechos. Es posible que la neurociencia pueda explicar nuestras capacidades morales e incluso nuestras creencias acerca de lo que es correcto o justo o bueno, pero no puede justificarlas. Como señala CELA CONDE, «los hallazgos neurocientíficos servirán para alcanzar un mayor conocimiento acerca de la naturaleza humana, pero

este no garantiza, por sí mismo, valores morales como puedan ser un mayor respeto a la vida, a la igualdad y a la libertad humanas» (CELA *et al.*, 2005: 132). De hecho, claramente la neuroética normativa presupone implícitamente un valor moral que da sentido a su propuesta: la evolución es moralmente valiosa. ¿Es adecuado este presupuesto? Creo que no: de este valor (sin más matizaciones) se seguiría también la corrección moral de conductas favorables para la evolución que no estaríamos dispuestos a aceptar. De manera que, aunque se admitiera que la evolución es un valor, no sería, desde luego, un valor absoluto.

Esto no quiere decir que no existan conexiones entre cómo funciona nuestro cerebro, la evolución y la moral. La neurociencia y la biología pueden enseñarnos mucho acerca de la naturaleza humana, y estos conocimientos son importantes para la construcción de una moral racional. No estoy sugiriendo fundamentar lo que debemos hacer en la naturaleza biológica del ser humano, pero sí verla como un límite que señala lo que una moral racional puede o no puede exigir. Consideremos el principio «debe implica puede». Se trata de una exigencia de racionalidad práctica, de acuerdo con la cual que haya una norma que obliga o prohíbe hacer algo requiere que lo ordenado pueda ser realizado o lo prohibido pueda ser omitido. Dado que la moral ideal debe ser concebida como un sistema racional de normas o principios, el principio «debe implica puede» es una condición de existencia de las normas morales[1]. Nuestra forma de pensar (es decir, nuestras capacidades determinadas por la estructura y el funcionamiento de nuestro cerebro), nuestra biología, nuestras emociones, etc. son datos que una ética debe tener en cuenta y que la neurociencia puede sacar a la luz. Lo neurológicamente imposible (como lo físicamente imposible) no puede ser ordenado, salvo que esa imposibilidad se pueda superar por medio de la educación. Y, entonces, la neurociencia puede aportar orientaciones eficaces para esa educación. Que el principio «debe entraña puede» pone lí-

[1] ¿Afecta este principio a la ley de Hume? Creo que puede decirse que no: si algo es imposible, de ahí se sigue que no existe una norma que lo prohíbe; esto implica que está permitido en sentido débil, no que exista una norma que lo permita (ALCHOURRÓN y BULYGIN). No estamos derivando, por tanto, una norma de un hecho, sino una proposición normativa acerca de la existencia o no de una norma de una proposición sobre la posibilidad de realizar determinada acción (VON WRIGHT, 1970; MORESCO, 2016).

mites a la moral es algo que reconocemos ya en ciertas situaciones: supongamos un caso particular en el que, de acuerdo con nuestros conocimientos acerca del funcionamiento del cerebro, un sujeto no ha actuado libremente en un momento determinado, esto es, no ha podido actuar de otra manera. Entonces no tenía la obligación de actuar de otra manera y no debe ser tenido por responsable. Por supuesto, salvo que haya alguna manera de superar ese impedimento, y entonces deberá tratar de superarlo. Pensemos también en el deber de ayudar a los países necesitados del que habla SINGER. O de la intuición de que causar un daño por medio de una acción suele ser más reprobable que permitir por omisión que ese mismo daño ocurra. Sería posible que la escasa eficacia de los deberes positivos generales se deba a factores evolutivos y neurofisiológicos. Si tales factores constituyeran una imposibilidad no superable, no podríamos afirmar la existencia de tales deberes. Si son superables, esto es, si no estamos determinados en este punto (y hay estudios sobre la plasticidad del cerebro que apuntan a la posibilidad de que esa determinación no sea tan rígida y, por tanto, que puede superarse por medio de la reflexión y la educación), entonces el deber existe y debemos esforzarnos por cumplirlo. Un sistema moral debe ser «neurofisiológicamente» plausible. Esto, sin embargo, no tiene por qué rebajar el nivel de exigencia moral: plausibilidad no es comodidad.

En definitiva, nuestras pasiones, intuiciones y emociones no son suficientes para fundamentar una ética universal, pero la neuroética puede proporcionar conocimientos que nos ayuden a extender de una manera artificial (es decir, cultural, por medio de la educación y, por tanto, bajo la vigilancia de la razón) las limitadas fronteras naturales de nuestra «simpatía», como proponía HUME. Conocer la naturaleza humana (y ser consciente de su dimensión biológica) no es solo una curiosidad teórica, sino también una necesidad práctica para construir nuestra dimensión moral.

Cambridge, agosto de 2016.

BIBLIOGRAFÍA

ALARCÓN CABRERA, C., 2001: «Valores y sentimientos: un enfoque freudoemotivista», *Anuario de Filosofía del Derecho,* núm. 18.

ALCHOURRÓN, C., 2010: «Fundamentos filosóficos de la lógica deóntica y la lógica de los condicionales derrotables», en MORESO, J. J., y RODRÍGUEZ, J. L. (eds.), *Fundamentos para una teoría general de los* deberes, Madrid, Marcial Pons.

ÁLVAREZ-DÍAZ, J. A., 2013: «Neuroética como neurociencia de la ética», en *Revista de neurología*, 57 (8).

AYALA, F., 2013: *Evolución, ética y religión*, Universidad de Deusto.

AYER, A. J., 1984: *Lenguaje, verdad y lógica,* Barcelona, Orbis.

BARROSO FERNÁNDEZ, O., 2014: «La tesis bio-antropológica del innatismo. Falacias descubiertas desde la reflexión filosófica acerca de la experiencia moral», en *Contrastes. revista internacional de Filosofía*, vol. XIX, núm. 2.

BARTA, R., 2013: *Cerebro y libertad. Ensayo sobre la moral, el juego y el determinismo*, México, FCE (versión electrónica).

BAYÓN, J. C., 2000: «Derrotabilidad, indeterminación del Derecho y positivismo jurídico», en *Isonomía*, núm. 13.

— 2001: «¿Por qué es derrotable el razonamiento jurídico?», *Doxa,* núm. 24.

BONETE PERALES, E., 2010: *Neuroética práctica. Una ética desde el cerebro*, Henao, Desclée de Brouwer.

BRAUN, R., 2014: «La neuroética y la naturalización de la filosofía moral», en *Persona*, núm. 17, Lima.

BUNGE, M., 2013: *Intuición y razón*, Buenos Aires, Debolsillo (versión electrónica).

CAMPS, V., 2013: «Neuronas y valores», en *Revista de Neurología*, núm. 57.

CASEBEER, W. D., 2003: «Moral cognition and its neural constituents», en *Nature Reviews/Neuroscience*, vol. 4.

CASEBEER, W. D., y CHURCHLAND, P., 2003: «The Neural Mechanisms of Moral Cognition: A Multiple-Aspect Approach to Moral Judgment and Decision-Making», en *Biology and Philosophy,* 18, Kluwer Academic Publishers.

CELA CONDE, C. J.; FERNÁNDEZ, A.; MARTY, G.; NADAL, M., y CAPÓ, M. A., 2005: «Derecho y neurociencia», en *Ludus Vitalis,* vol. XIII, num. 23.

CELA CONDE, C. J.; FLEXAS, A., y CHISTENSEN, J. F., 2009-2010: «Los valores humanos y el cerebro», en *Revista Taula. Quaderns de pensament*, núm. 42.

CORTINA, A., 2010: «Neuroética: ¿Las bases cerebrales de una ética universal con relevancia política?», en *Isegoría, Revista de Filosofía Moral y Política,* núm. 42.

— 2011: *Neuroética y neuropolítica. Sugerencias para la educación moral*, Madrid, Tecnos.

CHURCHLAND, P., 2013: «The impact of Neuroscience on Philosophy», *http://dx.doi.org/10.1016/j.neuron.*

— 2012: *El cerebro moral. Lo que la nerociencia nos cuenta sobre la moralidad,* Barcelona, Paidós (versión electrónica).

DAMASIO, A. R.: 2004, *El error de Descartes*, Barcelona, Crítica.

DEL BARCO COLLAZOS, J. L., 1984: «El *"is-ought passage"* y su evolución en el seno del neopositivismo», en *Thémata: Revista de filosofía,* núm. 1.

DIÉGUEZ, A., 1998: *Realismo científico. Una introducción al debate actual en la teoría de la ciencia*, Universidad de Málaga, 1998.

— 2002: «Realismo y epistemología evolucionista de los mecanismos cognitivos», en *Crítica*, vol. 34, núm. 102.

— 2003, «¿Qué es la epistemología evolucionista?», en *Teleskop*, vol. 1, núm, 3. DOI: *http://dx.doi.org/10.1016/S0896-6273(02)00763-8.*

DUSELL, E., 2001: «Algunas reflexiones sobre la "falacia naturalista" (¿Pueden tener contenidos normativos implícitos cierto tipo de juicios empíricos?)», en *DIÁNOIA*, año XLVI, núm. 46.

ESTANY, A., 2013: «La filosofía en el marco de las neurociencias», en *Revista de neurología*, núm. 56.

EVERS, K., 2010: *Neuroética. Cuando la maeria se despierta*, Buenos Aires-Madrid, Katz.

FLANAGAN, O., 1996: «Ethics Naturalized. Ethics as Human Ecology», en MAY, L.; FRIEDMAN, M., y CLARK, A., *Mind and Morals. Essays on Cognitiva Science and Ethics*, Cambridge, MA Mit Press.

FOOT, P., 1974: «Introducción», en *Teorías de la ética*, FCE, México.

— 1994, «El problema del aborto y la doctrina del doble efecto», en *Las virtudes y los vicios*, México, UNAM.

FRANKENA, W. K.: 1974, «La falacia naturalista», en FOOT, P., *Teorías de la ética,* México, FCE.

GARCÍA-ALBEA, J. E., 2011: «Usos y abusos de lo "neuro"», en *Revista de Neurología*, 52 (10), *www.neurologia.com.*

GARCÍA GARCÍA, E., 2014: «Neuropsicología del comportamiento moral», en DE LA TORRE DÍAZ, J., *Neurociencia, neuroética y bioética*, Universidad Pontificia de Comillas.

GARDERFÖRS, P., 2005: «Las intuiciones como conocimiento incompleto», en *Revista de Filosofía*, Universidad de Costa Rica, XLIII, 108.

GAZZANIGA, M. S., 2015: *El cerebro ético*, Barcelona, Paidós.

GÓMEZ MELO, J. D., y SERRANO-BOSQUET, F. J., 2015: «Neuroética fundamental. El intento actual del naturalismo ético por fundar una moral científica», en *Revista de Filosofía*, núm. 79, México.

GONZÁLEZ LAGIER, D., 1995: *Acción y norma en G. H. von Wright*, Madrid, Centro de Estudios Constitucionales.

— 2007: «Hechos y conceptos», en *Cuadernos electrónicos de Filosofía del Derecho*, núm. 15.

— 2009: *Emociones, responsabilidad y Derecho*, Madrid, Marcial Pons.

GREENE, J. D., 2012: «Del "es" neuronal al "debe" moral: ¿Cuáles son las implicaciones morales de la psicología moral neurocientífica?», en CORTINA, A. (ed), *Guía Comares de Neurofilosofía prácica*, Granada, Comares.

GREENE, J. D.; SOMMERVILLE, R. B.; NYSTROM, L. E.; DARLEY, J. M., y COHEN, J. D., 2001: «An fMRI Invesigation of Emocional Engagement in Moral Judgment», en *Science*, vol. 293 (*www.sciencemag.org*).

HAIDT, J., 2012: «El perro emocional y su cola racional: Un enfoque intuicionista social del juicio moral», en CORTINA A. (ed.), *Guía Comares de Neurofilosofía práctica*, Granada, Comares.

HAIDT, J., y CRAIG, J., 2004: «Intuitive ethics: how innately prepared intuitions generate culturally variable virtues», en *Dadedalus Fall*, American Academy of Arts & Sciences.

HARE, R. M., 1999: *Ordenando la ética. Una clasificación de las teorías éticas*, Barcelona, Ariel.

HAUSER, M. D., 2008: *La mente moral. Cómo la naturaleza ha desarrollado nuestro sentido del bien y del mal*, Barcelona, Paidós.

HIERRO, J., 1970: *Problemas del análisis del lenguaje moral*, Madrid, Tecnos.

HUDSON, W. D., 1987: *La filosofía moral contemporánea*, Madrid, Alianza Editorial.

HUME, D.: 1981, *Tratado de la naturaleza humana*, Madrid, Orbis.

— 1993, *Investigación sobre los principios de la moral*, Madrid, Alianza Editorial.

IACOBONI, M., 2013: *Las neuronas espejo. Empatía, neuropolítica, autismo, imitación o de cómo entendemos a los otros*, Katz.

KAHNEMAN, D., 2012: *Pensar rápido, pensar despacio*, Barcelona, Mondadori (versión electrónica).

KANT, I., 1973: *Fundamentación de la metafísica de las costumbres*, Ed. Espasa Calpe, Madrid.

— 1993: *Crítica de la razón pura*, Barcelona, Círculo de lectores.

LAKOFF, G., y JOHNSON, M., 1995: *Metáforas de la vida cotidiana*, Cátedra, col. Teorema.

LARIGUET, G., 2014: «Ética, giro experimentalista y naturalismo débil», en *Estudios de Filosofía Práctica e Historia de las Ideas*, vol. 16, núm. 2.

LUQUE, P., 2011: «Las dos vías del realismo moral», en *Discusiones*, núm. X.

MACINTYRE, A., 1959: «Hume on "is" and "ought"», en *The Philosophical Review,* vol. LXVIII, núm. 4.

MANYA, F., y GOMES, C., 2003: «Técnicas "de resolución" de problemas de satisfacción de restricciones», en *Inteligencia Artificial, Revista Iberoamericana de Inteligencia Artificial*, núm. 19.

MARTÍNEZ, M., 2003: «La falacia naturalista y el argumento de la pregunta abierta», *Universitas Philosophica,* 40-41, Bogotá.

MOORE, G. E. M., 1959: *Principia Ethica*, México, UNAM.

MORA, F., 2007: *Neurocultura*, Madrid, Alianza.

MORESO, J. J., 2016: «Von Wright sobre "ad impossibilia nemo tenetur"», en *Doxa. Cuadernos de Filosofía del Derecho,* núm. 39, en prensa.

MORESO, J. J., y RODRÍGUEZ, J. L., 2010: «Carlos Alchourrón y la máxima de mutilación mínima», en ALCHOURRÓN, C., *Fundamentos para una teoría general de los deberes*, Madrid, Marcial Pons.

MOULINES, U., 1991: «Hechos y valores: Falacias y metafalacias. Un ejercicio integracionista», en *Isegoría*, núm. 3.

NAVARRO, P., 2007: «Normas condicionales y falacia naturalista», en *DOXA, Cuadernos de Filosofía del Derecho*, núm. 30.

OSEGUERRA, J., 2013: «Los argumentos antinaturalistas de Moore», en *Revista de Estudios Filosóficos*, núm. 1.

OVEJERO, F., 2012: «La naturalización de la teoría social. Diez tesis», en *Encrucijadas. Revista crítica de ciencias sociales*, núm. 4.

PALLARÉS DOMÍNGUEZ, D. V., 2013: «Críticas y orientaciones para el estudio en neuroética», en CONILL SANCHO, J., y PÉREZ ZAFRILLA, P. J., *Retos actuales de la neuroética, Recerca. Revista de pensament i anàlisi,* núm. 13.

PÉREZ BERMEJO, J. M., 1999: «Simpatía y motivación. Las oscilaciones de la moral simpatética entre el sentido y la sensibilidad», en *Telos*, vol. VIII, núm. 1.

PÉREZ ZAFRILLA, P. J., 2013: «Implicaciones normativas de la psicología moral. Jonathan Haidt y el desconcierto moral», en *Δαι΄μων. Revista Internacional de Filosofía,* núm. 59.

PIGDEN, Ch. R., 2004: «El naturalismo», en SINGER, P., *Compendio de ética*, Madrid, Alianza Editorial.

PRICHARD, H. A., 1999: «Does Moral Philosophy Rest on a Mistake?», en *Moral obligation. Essays and Lecturers*, Oxford Clarendon Press.

PUTNAM, H., 2007: «¿Por qué la razón no puede ser naturalizada?», en *Signos Filosóficos*, vol. IX, núm. 18.

QUINE, W. V., 1970: «Natural Kinds», en RESCHER, N. *et al., Essays in Honor of Carl G. Hempel*, Dordrecht, D. Reidel.

RAWLS, J., 2006: *Teoría de la justicia*, México, Fondo de Cultura Económica.

RIVERA LÓPEZ, E., 2002: *Naturalismos Éticos*, en PEREZ, D. (comp.), *Los Caminos del Naturalismo: Mente, conocimiento y moral*, Buenos Aires, EUDEBA.

RIZZOLATI, G., y CRAIGHERO, L., 2004: «The Mirror Neuron-System», *Annuary Review of Neuroscience,* 27, doi: *10.1146/annurev.neuro.27.070203.144230.*

ROSKIES, A., 2002: «Neuroethics for the New Millenium» en *Neuron*, vol. 35, 1.

ROSS, A., 1963: *Sobre el Derecho y la Justicia*, Buenos Aires, EUDEBA.

ROSS, W. D., 1994: *Lo correcto y lo bueno*, Salamanca, Sígueme.

RUSE, M., 2004: «La significación de la evolución», en SINGER, P., *Compendio de ética*, Madrid, Alianza Editorial.

SALLES, A., 2013: «On the Normative Implications of Social Neuroscience», en CONILL SANCHO, J., y PÉREZ ZAFRILLA, P. J., *Retos actuales de la neroética, Recerca. Revista de pensament i anàlisi*, núm. 13.

SEARLE, J., 1980: *Actos de habla*, Madrid, Cátedra.

SHAND, B., 2014: «Neuronas espejo y simpatía en Adam Smith: comparación de dos perspectivas sobre la empatía, frente al reduccionismo científico», *Revista Colombiana de Filosofía de la Ciencia,* 14.29.

SINGER, P., 2012: «Ética e intuiciones», en CORTINA, A., *Guía Comares de Neurofilosofía práctica*, Granada, Comares.

STEVENSON, Ch., 1993: «El significado emotivo de los juicios éticos», en AYER, J., *El positivismo lógico*, México, Fondo de Cultura Económica.

TOVAR, J., 2008: «El papel de las emociones y de la razón en el juicio moral», en *Saga. Revista de estudiantes de Filosofía*, Bogotá, Universidad Nacional de Colombia.

VON WRIGHT, G. H., 1963: *The Varieties of Goodness*, Routledge & Kegan Paul. Existe traducción al castellano, *La diversidad de lo bueno,* Marcial Pons, 2010 (trad. de Victoria Roca Pérez y Daniel González Lagier).

— 1970, *Norma y acción. Una investigación lógica*, Madrid, Tecnos.
— 1976, *Un ensayo de lógica deóntica y la teoría general de la acción*, México, UNAM.
— 1983, «The Foundation of Norms and Normative Statements», en *Practical Reason*, Basil Blackwell.
— 1985: «Is and Ought», en BULYGIN, E.; GARDIES J. L., y NIINILUOTO, I., *Man, Law and Moderns Forms of Life*, D. Reidel.
— 2001, «Valorar, o cómo hablar de lo que se debe callar. Nuevas bases para el emotivismo moral», en *Anuario de Filosofía del Derecho*, núm. 18.
WIDOW LIRA, F., 2015: «La ley de Hume en Hume. La discusión de la interpretación analítica de Treatise III, 1, i», en *Anales del Seminario de Historia de la Filosofía*, vol. 32, núm. 2.
ZAVADIVKER, N., 2008: «Alfred Ayer y y la teoría emotivista de los enunciados morales», en *Anuario Filosófico*, XLI/3.
— 2014: «Enre el naturalismo y la normatividad: los dilemas bioéticos bajo el *scanner*», en *Cuadernos de Neuropsicología*, núm. 1.

COLECCIÓN «FILOSOFÍA Y DERECHO»

ÚLTIMOS TÍTULOS PUBLICADOS

(Véase la lista completa en *www.filosofiayderecho.es/titulospublicados.php*)

Positivismo jurídico lógico-incluyente
Juliano Maranhão

¿Las normas lógicamente derivadas de normas válidas son válidas? A partir de esta cuestión, en el centro de la tensión entre cognición y volición como fundamento de validez del derecho, el autor busca habilitar una alternativa al positivismo jurídico ignorada por el debate contemporáneo en la teoría analítica del derecho: el positivismo jurídico lógico-incluyente, según el cual las razones jurídicas vinculantes son aquellas derivadas lógicamente del contenido de actos de voluntad dotados de autoridad. La obra mantiene un debate sobre lo que serían las tesis mínimas de una teoría positivista, que lleva al autor a defender la indisociabilidad entre una teoría de la interpretación y una teoría de la validez de las normas jurídicas. Con ello, las derivaciones lógicas si no producen normas válidas pertenecientes al ordenamiento con el mismo *estatus ontológico* que las normas explícitas, al menos constituyen razones jurídicas vinculantes, en la medida en que, en la actividad interpretativa, la producción de normas jurídicas se concibe como actividad racional.

De este modo, la interpretación se entiende como una actividad de explicación del contenido de las normas, a partir de las inferencias comunes a la comunidad jurídica, bajo el presupuesto de una legislación racional. Esto hace que no solo las deducciones del contenido de las reglas sean razones vinculantes sino también las inducciones y las abducciones (inferencia a la mejor explicación posible), que pueden llevar a principios o valores como parte de las razones jurídicas derivadas de las fuentes dotadas de autoridad. Pero se invierte el orden de justificación: no son los principios o valores los que otorgan fuerza vinculante a las reglas, son las reglas, dotadas de fuente, que confieren fuerza a los principios o valores como razones vinculantes dentro de una determinada teoría interpretativa, solo y en la medida en que puedan servir de explicación plausible de su contenido.

Al trazar distinciones con relación a teorías divergentes como el iusnaturalismo, el derecho como integridad, el positivismo excluyente y el positivismo incluyente, el autor sitúa el positivismo lógico-incluyente como alternativa a una metodología jurídica valorativa, pero moralmente neutra.

Democracia participativa epistémica
Sebastián Linares

Democracia participativa epistémica es un libro que puede leerse en dos niveles: filosófico y de ingenieria constitucional. En el nivel filosófico, el libro procura justificar las bases de una teoría de la democracia participativa y no populista. La obra argumenta que los mejores ideales de democracia (que son deliberativos y epistémicos) necesariamente desembocan en una teoría aplicada de la democracia fuertemente comprometida con el derecho a participar en pie de igualdad, pero bajo un sistema institucional dirigido a incentivar de manera estructural la generación autonóma de preferencias ciudadanas informadas y orientadas al bien común. En el nivel de ingeniería constitucional el libro intenta articular con detalle un elenco de instituciones novedosas que persiguen el cometido epistémico modesto de alcanzar decisiones

suficientemente buenas en el largo plazo. Entre esas instituciones está un nuevo régimen de voto ciudadano, una nueva regulación de las abstenciones, la institución de la promesa pública del votante y un nuevo sistema de democracia directa activado por la ciudadanía desde abajo.

La obra desarrolla una crítica reflexiva de las teorías dominantes contemporáneas de la legitimidad, entre ellas la teoría pluralista de la democracia de Robert Dahl, la teoría agonista de la democracia de Ernesto Laclau, la concepción elitista de la democracia de Joseph Schumpeter o Giovanni Sartori, e incluso de las concepciones participativas puramente agregativas de la democracia.

Stare indecisis
Las alteraciones del precedente en la Corte Suprema de Estados Unidos, 1946-1992
Saul Brenner y Harold J. Spaeth

Aunque el concepto de precedente es básico para que opere el sistema jurídico, todavía no se han realizado estudios empíricos comprehensivos acerca de por qué los jueces de la Corte Suprema de los Estados Unidos eligen alterar un precedente. Este libro intenta llenar aquel vacío analizando las decisiones que alteraron formalmente precedentes en las Cortes de Vinson, Warren y Burger, así como en los seis primeros ejercicios de la Corte de Rehnquist —un periodo de cuarenta y seis años (1946-1992)—. Los autores resumen estudios previos sobre el precedente y la Corte, evalúan la votación en asamble de los jueces y compilan una lista de casos que rechazan precedentes y de los casos rechazados. Adicionalmente, trazan una distinción entre stare decisis personal e institucional. Utilizando el modelo actitudinal de toma de decisiones de la Corte Suprema, que normalmente es visto como antitético al modelo jurídico de votación, los autores descubren que son las ideologías individuales de los jueces las que explican su comportamiento al votar.

El iusnaturalismo egoísta de Thomas Hobbes
Carlos Isler Soto

El iusnaturalismo egoísta de Thomas Hobbes es una interpretación general de la teoría jurídica de Hobbes, y muestra la dependencia de dicha teoría respecto de la filosofía moral y epistemológica del filósofo inglés. Hobbes es un iusnaturalista, en cuanto cree que el cumplimiento de un nivel mínimo de moralidad es condición necesaria de juridicidad de una norma jurídica. Dicho nivel mínimo de moralidad corresponde a la satisfacción de los intereses más básicos del sujeto individual y ello, junto a ciertas premisas de carácter antropológico —como el carácter constitutivo del mundo por parte del lenguaje y el egoísmo sicológico— llevarán a Hobbes a concluir que la existencia y el contenido de los ordenamientos jurídicos es relativa a los sujetos individuales. De este modo, asimismo, toda obligación jurídica será un caso de obligación moral.

Se explica, en conexión con el tema anterior, el carácter normativo de la teoría jurídica hobbesiana, en cuanto el filósofo de Malmesbury la desarrolla teniendo en vista lo que cree ser el bien para el hombre, y se muestra que Hobbes anticipa discusiones y doctrinas propias de autores contemporáneos como Hart, Raz o Finnis, y puede, de este modo, ser un importante referente en la discusión actual en filosofía del derecho.

Identificación y justificación del derecho
Segunda edición
Josep M. Vilajosana

Este libro supone una notable contribución a la reflexión clara, ordenada y coherente acerca de algunos de los problemas más relevantes de la filosofía del derecho.

La primera parte gira en torno a los problemas de identificación del derecho, concretados en las siguientes preguntas: ¿Cuándo existe el derecho en una determinada sociedad? ¿Está el derecho relacionado con la moral? ¿Está el derecho determinado? La segunda parte versa acerca de los problemas de justificación tanto de la obediencia al derecho como de la imposición de penas y la imposición jurídica de la moral.

El tratamiento de estas cuestiones se hace no con una vocación exhaustiva, sino selectiva. No importa tanto la reconstrucción completa de las doctrinas de los autores más importantes, sino la exposición crítica de los principales argumentos esgrimidos a la hora de abordar los citados problemas. A través de la comprensión de tales argumentos se pretende que el lector pueda formarse su propia opinión acerca de estas cuestiones.

Estas razones hacen que la presente obra sea especialmente recomendable como manual de filosofía del derecho.

Suerte moral, castigo y comunidad

Un análisis de la relevancia moral de la suerte en el resultado

Gustavo A. Beade

La suerte juega un papel importante en la determinación de nuestros juicios de responsabilidad y nos obliga, a menudo, a reflexionar sobre nuestras nociones de moralidad. Por ejemplo, nos preguntamos por qué un asesino debe ser castigado más severamente que quien solo intentó matar a otra persona, y también por qué damos más reconocimiento a quien descubre una vacuna que permite curar una enfermedad que a quien solo intentó hacerlo. Si bien algunos filósofos y teóricos del derecho admiten esta circunstancia, existen muchos otros que niegan la posibilidad de que la responsabilidad pueda apoyarse en circunstancias que están fuera de nuestro control.

El objetivo del libro es mostrar que nuestras comunidades reconocen la existencia de la suerte y la caracterizan como aquello que se encuentra más allá de nuestra voluntad y nuestras intrincadas planificaciones. Si aceptamos este punto de partida, es posible también asumir la presencia en nuestras vidas de la suerte moral: ser reconocido o castigado en parte por circunstancias que están fuera de nuestro control. Este trabajo pretende defender la influencia de la suerte en nuestros juicios de responsabilidad moral y legal y también señalar sus consecuencias en el modo en el que fundamentamos el merecimiento, el reconocimiento, la inculpación y el castigo.

FILOSOFÍA Y DERECHO

A la sombra de Hume

Un balance crítico del intento de la neuroética de fundamentar la moral

¿Son nuestras opiniones morales fruto de nuestros razonamientos o, por el contrario, son el resultado de emociones y sentimientos, como pensaba Hume? Los espectaculares avances de la neurociencia en el estudio del funcionamiento del cerebro humano han propiciado propuestas de entender a la ética como un conjunto de intuiciones o emociones vinculadas a la evolución humana, cuya función es garantizar la supervivencia de la especie. Algunos autores no solo tratan de explicar de esta manera la capacidad de comportarse moralmente de los seres humanos, sino que defienden la existencia de una ética universal, basada en tales intuiciones y en el funcionamiento del cerebro, que debería servir de fundamento para nuestros códigos morales. Estas concepciones podrían verse como parte de un proceso de «naturalización» de la filosofía práctica. En este libro se pretende someter a examen algunas de las aportaciones de la llamada neuroética, sosteniendo que muchas de ellas incurren realmente en argumentaciones muy débiles, la mayoría de las veces por falta de rigor conceptual y filosófico.

Daniel González Lagier

Catedrático de Filosofía del Derecho de la Universidad de Alicante, donde imparte diversas asignaturas, entre las que destacan las de Lógica y Argumentación y la de Argumentación en Materia de Hechos del Máster de Argumentación Jurídica de dicha Universidad. Es también profesor del Máster en Razonamiento Probatorio de la Universidad de Girona. Es autor de los libros *Acción y norma en G. H. von Wright*, *G. H. von Wright y los conceptos básicos del derecho*, *Las paradojas de la acción*, *Quaestio Facti. Ensayos sobre pruebas, causalidad y acción*, *Emociones, responsabilidad y derecho*, y numerosos capítulos de libros y artículos.

ISBN: 978-84-9123-421-0

9 788491 234210

CEA APQ Calidad en Edición Académica Academic Publishing Quality

Marcial Pons